THE UNIVERSITY *of* EDINBURGH

Effects of Metal Ions on the Structural and Biochemical Properties of Trypanosomatid Phosphoglycerate Mutases

Fazia Adyani Ahmad Fuad

Thesis submitted for the degree of

Doctor of Philosophy

Structural Biochemistry Group
Institute of Structural and Molecular Biology
School of Biological Sciences
University of Edinburgh
Scotland
United Kingdom
May 2012

Abstract

Flagellate protozoa from the order Trypanosomatida have developed a range of strategies to survive in their mammalian hosts. A consequence is that the glycolytic pathway has assumed an important role, especially in bloodstream-form *Trypanosoma brucei*, where it is essential as the sole producer of ATP. The seventh enzyme in the pathway, 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (iPGAM) is particularly attractive as a drug target because it shares no common properties with the corresponding enzyme in humans. This enzyme catalyses the conversion of 3PGA to 2PGA, with the requirement for metal ions to assist the catalytic function. In this study, two important biochemical and structural aspects of the enzyme were investigated: i) The *in vitro* and *in vivo* requirements for biologically relevant metal ions to support the activity of iPGAM, and ii) The ability of trypanosomatid iPGAM to exist in multiple conformations and oligomeric states in solution.

The maximum activity of iPGAM *in vitro* requires Co^{2+}, but this cannot be the case *in vivo* where ICP-OES analyses confirmed that Co^{2+} was essentially undetectable in *T. brucei* cytosolic fractions. The activity of iPGAM *in vivo* is therefore one of the lowest among the glycolytic enzymes. By contrast, Mg^{2+} and Zn^{2+} were found to be the most abundant metals in both cytosolic fractions and in purified bacterially expressed iPGAM. Our newly-developed multimode-plate reader discontinuous assay further revealed that of the biologically relevant metals, only Mg^{2+} can support iPGAM activity, but at less than 50% of the level of Co^{2+}. By contrast, Zn^{2+} strongly inhibits iPGAM. This assay which was developed with minimal metal interference on the coupling enzymes, also showed that in solution, the ratio of the concentrations of 3PGA:2PGA (substrate:product) at equilibrium is not 1:1 as observed in the crystal structure, but is in fact 12:1, which may be due to the tighter binding of 2PGA to the enzyme.

A series of biophysical analyses, notably by SEC-MALS showed that iPGAM from *Leishmania mexicana*, another trypanosomatid protozoan parasite exists in different forms and oligomeric states in solution, either as the closed-form monomer, open-

form monomer, or closed/open-form dimer which can be successfully separated by ion-exchange chromatography. The open-form *Lm*iPGAM is particularly relevant for drug development, as the catalytic site in the closed-form structure is poorly inaccessible. Both virtual and high-throughput screening approaches were used to identify novel potential inhibitors. Out of a collection of 11 compounds tested at 1 mM, two showed substantial inhibition with 49% and 14% remaining activity. Taken together, the findings from this study demonstrated the potential of iPGAM to be a key modulator in controlling glycolytic flux in trypanosomes, and thus further validated it as an important drug target.

Declaration

I hereby declare that the research recorded in this thesis and the thesis itself was composed and originated entirely by myself in the School of Biological Sciences at The University of Edinburgh, United Kingdom, except when otherwise stated.

Fazia Adyani Ahmad Fuad
Edinburgh, May 2012

Acknowledgements

First and foremost, I would like to express my special note of gratitude to my supervisors, Professor Malcolm Walkinshaw and Dr. Linda Gilmore for their endless guidance and support throughout my PhD years. It has been a wonderful opportunity to be part of their research group, and I am earnestly grateful for their excellent supervision and mentoring commitments.

Specifically for the researchers in the trypanosomatid glycolytic enzymes research group, Dr. Hugh Morgan, Dr. Matt Nowicki and Dr. Iain McNae, thank you very much for the bright ideas, guidance and the wonderful discussions. Also, warm thanks are noted to the Centre for Translational and Chemical Biology (CTCB), especially to Mrs Sandra Bruce, Dr. Martin Wear, Dr. Janice Bramham, Dr. Elizabeth Blackburn and Dr. Paul Taylor for their technical help and assistance throughout the period of study. Not forgetting all the current members of the Walkinshaw group in Swann Level 3 and also the former ones, many thanks for sharing ideas, offering help and comments during my PhD years.

A special thank-you note is forwarded to Dr. Marjorie Harding for the wonderful discussions about metals in protein structures, Dr. Lorna Eades for the metal analyses, Dr. Andrew Cronshaw for assisting the mass-spectrometry experiments and Dr. Douglas Houston for performing part of the virtual screening analyses. My sincere gratitude also goes to the NIH research group in Bethesda, USA for the high-throughput screens and also the Brussels research group, especially Professor Paul Michels for the various biological samples, the kind assistance and helpful discussions. This work has been made possible with collaborations and guidance from all.

Last but certainly not least, to a very special person in my life, my beloved husband Mohd Zulhakimi Ab Razak, thank you very much for the love, guidance and not forgetting the academic discussions throughout our lives together and PhD studies. My very special gratitude also goes to my beloved parents and family for the

encouragement in my academic studies, love, care and never-ending support, throughout very many years. Also, to all my teachers and close friends, a big thank you for the encouragement and kind support.

These PhD studies were also made possible with the generous scholarship and funding that have been awarded by the government of Malaysia. For this, an appreciation is addressed to the Ministry of Higher Education Malaysia and Universiti Sains Malaysia.

Abbreviations

2PGA	2-phosphoglycerate
3PGA	3-phosphoglycerate
ADP	Adenosine Diphosphate
ALD	Aldolase
APS	Ammonium Persulphate
ATP	Adenosine Triphosphate
BaiPGAM	*B. anthracis* 2,3-bisphosphoglycerate-independent phosphoglycerate mutase
BSA	Bovine Serum Albumin
BsiPGAM	*B. stearothermophilus* 2,3-bisphosphoglycerate-independent Phosphoglycerate Mutase
CRC	Concentration Response Curves
CTCB	Centre for Translational and Chemical Biology
CV	Column Volumes
DLS	Dynamic Light Scattering
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dPGAM	Cofactor-dependent Phosphoglycerate Mutase
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic Acid
EDULISS	EDinburgh University LIgand Selection System
ENO	Enolase
G6DPH	Glucose-6-phosphate dehydrogenase
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
His-tagged	Histidine-tagged
HXK	Hexokinase
ICP-MS	Inductively Coupled Plasma-Mass Spectrometry

ICP-OES	Inductively Coupled Plasma-Optical Emission Spectrometry
IMAC	Immobilised Metal Affinity Chromatography
iPGAM	Cofactor-independent phosphoglycerate mutase
IPTG	Isopropyl β-D-1-thiogalactopyranoside
ITC	Isothermal Titration Calorimetry
K_{av}	Gel phase distribution coefficient
K_d	Dissociation Constant
K_m	Michaelis Constant
LIDAEUS	Ligand Discovery at Edinburgh UniverSity
LDH	Lactate Dehydrogenase
*Lm*iPGAM	*L. mexicana* 2,3-bisphosphoglycerate-independent phosphoglycerate iPGAM
LOD	Limit Of Detection
LOPAC	Library Of Pharmacologically Active Compounds
MALDI-TOF	Matrix Assisted Laser Desorption Ionisation- Time Of Flight
MCS	Multiple Cloning Site
MESPEUS	Metal Sites in Proteins at Edinburgh UniverSity
MLogP	Moriguchi Water / Octanol Partition Co-efficient
NAD^+	Nicotinamide Adenine Dinucleotide
NADH	Reduced Nicotinamide Adenine Dinucleotide
NCBI	National Center for Biotechnology Information
NIH	National Institutes of Health
PCR	Polymerase Chain Reaction
PEP	Phosphoenolpyruvate
PFK	Phosphofructokinase
PGI	Phosphoglucose Isomerase
PGK	Phosphoglycerate Kinase

PYK	Pyruvate Kinase
qHTS	quantitative High-Throughput Screening
RF	Radio Frequency
R_g	Radius of Gyration
R_h	Hydrodynamic Radius
RNA	Ribonucleic Acid
RNAi	RNA interference
SAR	Structure Activity Relationship
SDS-PAGE	Sodium Dodecyl Sulphate-Polyacrylamide Gel Electrophoresis
SEC-MALS	Size Exclusion Chromatography-Multi-angle Light Scattering
SOD	Superoxide Dismutase
SPR	Surface Plasma Resonance
STP	Surface Triplet Propensities
*Tb*iPGAM	*T. brucei* 2,3-bisphosphoglycerate-independent Phosphoglycerate Mutase
TCA	Trichloroacetic Acid
TEA-HCl	Triethanolamine-Hydrochloride
TEMED	Tetramethylethylenediamine
TIM	Triosephosphate isomerase
USFRAT	Ultra-Fast Shape Recognition with Atom Types
WHO	World Health Organisation

Contents

Chapter 1:

Introduction

1.1 Leishmaniasis and African trypanosomiasis: The causes of the diseases and drug development at present

For centuries, a group of kinetoplastid parasites, *Leishmania spp* and *Trypanosoma brucei* has been and still is one of the biggest threats to the inhabitants of many parts of the world (Steverding 2008; von Geldern et al. 2011). Both parasites caused fatal neglected diseases: leishmaniasis and human African trypanosomiasis respectively, resulting in epidemically affected 88 countries worldwide (Singh 2006). Tragically, the main factor contributing to the emergence of the diseases in the infected areas is poverty, which is associated with a number of social aspects: malnutrition, displacement, poor housing, illiteracy, weakness of the immune system and lack of resources (http://www.who.int/leishmaniasis/burden/en/). Since effective vaccines against the diseases are still under development, chemotherapy is the only mode of control (Chawla and Madhubala 2010). Advances in pharmacology have led to the introduction into the market of a number of drugs, yet unfortunately, these have promoted toxic side effects in the treated patients, have required long-term treatments and have generated resistance mechanisms (Castillo et al. 2010). Hence, there is urgency in discovering new anti-kinetoplastid drugs for the treatment of these diseases.

The severity of leishmaniasis is exhibited by the number of people developing the symptomatic disease: cutaneous (1-1.5 million) and visceral (0.5 million), resulting in around 70 000 deaths per year (Murray et al. 2005). Third-world countries experienced the most infections, where 90% of both cutaneous and visceral leishmaniases have developed in some parts of Asia, Africa and Southern America. *Leishmania* spp are classified under the kingdom Protista and phylum Euglenozoa, and are mainly characterised by the presence of an associated cytostome (mouth) supported by one of three microtubule groups that arise from flagellar bases (Singh 2006). The vector for the parasites is the phlebotomine sandfly, which transmits the parasite into

the mammalian host. The life cycle of the parasite generally occurs with the existence of two morphologically distinct forms: a motile flagellated form (promastigotes) and an intracellular non-flagellated form (amastigotes) which exist in the sand fly and human stages, respectively (Castillo et al. 2010). Four main clinical syndromes are identified depending on the symptoms of the disease: cutaneous (Minodier and Parola 2007), muco-cutaneous (Amato et al. 2008), visceral (Sundar and Chatterje 2006) and post-kala-azar dermal leishmaniasis (Zijlstra et al. 2003). At present, anti-leishmanial control depends solely on chemotherapy, with several drugs available for treatment: pentavalent antimonials, amphotericin and more recently, paromomycin, miltefosine and liposomal amphotericin B (Castillo et al. 2010). Unfortunately these drugs have limitations in their capability to treat the disease (Table 1.1). Hence, the present scenario exhibits the urgency of discovering more affordable, less toxic, orally-taken drugs that are efficacious in HIV patients and against resistant strains (Golgher et al. 2011).

Human African trypanosomiasis, also known as sleeping sickness, is another severe disease, and is caused by the protozoan parasite *T. brucei*, affecting the sub-Saharan Africa regions with 50 000-70 000 cases currently reported (Brun et al. 2010). The disease remains endemic in the region due in part to the suitability of the habitats for the transmitting vector, the tsetse flies of the genus *Glossina* (Brun et al. 2010). The parasite causes different variants of trypanosomiasis caused by different sub-species: i) *T. brucei gambiense*, ii) *T. brucei rhodesiense* (Barrett et al. 2003) and iii) *T. brucei brucei* (Vanhamme et al. 2003). Out of the three sub-species, only *T. brucei brucei* is not human infective, and thus has been used for laboratory research. On the other hand, the other two subspecies: *T. brucei gambiense* and *T. brucei rhodesiense* cause a slowly-developed chronic disease and an acute form of the disease, respectively. The life cycle of the parasite involves stages in human and the tsetse fly which mainly exist as: i) the non-pathogenic form known as procyclic trypomastigotes which refers to the parasite that lives in the midgut of the fly and ii) bloodstream trypomastigotes that are associated with the pathogenic form which lives in human blood (Silverman and Bangs 2012). The current treatment depends on the disease stage which can be differentiated

into two major phases: a) the first infectious stage and b) the second neurological stage when the parasite has crossed the blood-brain barrier and affects the central nervous system (http://www.who.int/mediacentre/factsheets/fs259/en/). The treatment also depends on the causative agent (Brun et al. 2010), for example pentamidine, eflornithine and melarsoprol for *T. brucei gambiense*, and suramin and melarsoprol for *T. brucei rhodesiense* (Table 1.1). More recently, nitroimidazoles (antibacterial and antiprotozoal drugs) (Trunz et al. 2011), as well as SCYX-7158, an orally-active benzoxaborole (Jacobs et al. 2011; Jacobs et al. 2011) have been proposed to be promising leads for the disease, including at the second stage. While the present drugs are available for treatment, there are substantial problems of toxicity, and thus new drugs are urgently needed as safer alternatives.

Table 1.1 The current treatments for human African trypanosomiasis and leishmaniasis. Most information was obtained from the WHO website, (http://www.who.int/mediacentre/factsheets/fs259/en/) and Castillo et al. 2010. The table was adapted from a presentation by Dr. Linda Gilmore.

Drug	Date first used	Limitations
Human African Trypanosomiasis		
Suramin (First stage treatment)	1921	Toxicity, only effective in the early stage, injected
Pentamidine (First stage treatment)	1941	Toxicity, only effective in the early stage, required injection and resistance
Melarsoprol (Second stage treatment)	1949	Toxicity, resistance and required injection
Eflornithine (Second stage treatment)	1990	Costly, injected and only effective against *T. gambiense*
Combination of eflornithine and nifurtimox (Second stage treatment)	2009	Not effective against *T. rhodesiense*.
Leishmaniasis		
Pentamidine	1939	Toxicity, resistance and required injection
Pentavalent antimonials	1950	Toxicity, resistance and required injection
Liposomal amphotericin B	1990	Costly and required injection
Miltefosine	2002	Contraindicated in pregnancy
Paromomycin	Available in 1960 but abandoned until 2005	Toxicity and required injection
Fexinidazole	2010 (predicted to be a safe and effective oral drug candidate)	Not identified yet

Various aspects of parasite metabolism have been studied, in order to discover new drugs for the treatment for the diseases. For instance, enzymes in the glycolytic pathway such as phosphofructokinase (PFK) (EC 2.7.1.11), phosphoglycerate kinase (PGK) (EC 2.7.2.3), cofactor-independent phosphoglycerate mutase (iPGAM) (EC 5.4.2.1), enolase (ENO) (EC 4.2.1.11) and pyruvate kinase (PYK) (EC 2.7.1.40) (Verlinde et al. 2001; Albert et al. 2005; Coley et al. 2011), the pentose phosphate pathway (glucose-6-phosphate dehydrogenase (G6PDH) (Shreedhara et al. 2011) and folate metabolism (Gilbert 2002) as well as thiolases (Sommer and Wang 1994) have been validated as drug targets. For these enzymes, the knock-out of the corresponding target genes by RNA interference experiments (RNAi) (Albert et al. 2005; Owino et al. 2008) resulted in the death of the cultured parasites, which provides hope for new drug discoveries.

1.2 Targeting enzymes involved in glycolysis for trypanosomatid chemotheraphy

The role of glycolysis in nearly all living organisms has been elucidated for many years, indicating that it is one of the most primeval metabolic pathways that has ever existed (Fothergill-Gilmore and Michels 1993; Romano and Conway 1996). The glycolytic pathway essentially is a key player in converting glucose to pyruvate through a series of biochemical reactions, producing high-energy compounds such as adenosine triphosphate (ATP) and reduced nicotinamide adenine dinucleotide (NADH). In trypanosomatids, glycolysis is indeed very important, and in particular, plays an essential role as the sole producer of ATP for the bloodstream form of *T. brucei*. While glycolytic enzymes in humans and most other organisms are located in the cytosol, trypanosomatids possess specialised peroxisome-like organelles called glycocomes (Figure 1.1), which sequester the first seven enzymes in the pathway (Opperdoes and Borst 1977; Verlinde et al. 2001; Gualdrón-López et al. 2011). In this case, the initial glucose molecule is converted to 3-phosphoglycerate (3PGA) in the glycosome, before being transported through the membrane into the cytosol, where the other three enzymes, iPGAM, ENO and PYK, are located (Verlinde et al. 2001) (Figure 1.2). For

over ten years, several enzymes in the glycolytic pathway of *T. brucei* have been validated as potential drug targets, where any inhibition of the enzymes involved should lead to the disruption of the glycolytic flux, thus resulting in the death of the parasite. The majority of these enzymes exhibit distinct molecular properties from the corresponding enzymes in humans, offering possibilities for discovering parasite-specific potent inhibitors for drug development (Verlinde et al. 2001).

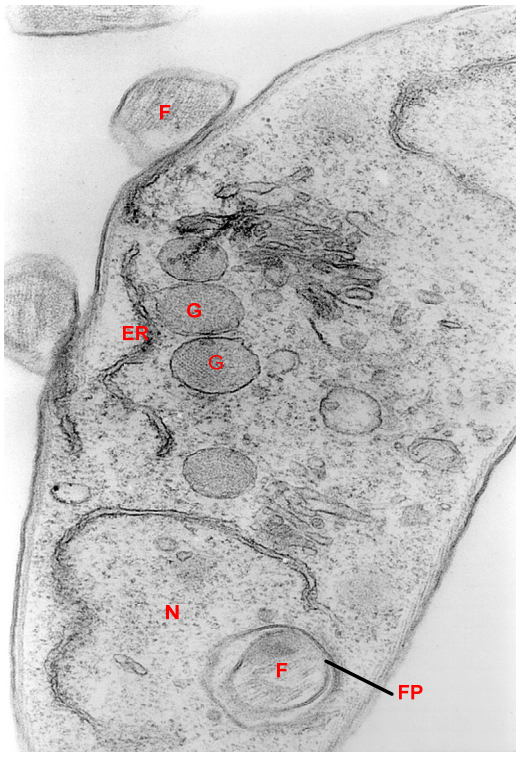


Figure 1.1 Electron micrograph of bloodstream-form *T. brucei* indicating the presence of various organelles in the parasite cell, which are labelled: G (glycosome) where the majority of glycolytic enzymes are located, ER (endoplasmic reticulum), F (flagellum), FP (flagellar pocket) and N (nucleus). Precise identification of the organelles is not straight-forward and requires immunostaining to identify markers and/or serial sectioning. iPGAM however, is present in the cytosol. The diagram was obtained from Prof. Paul Michels.

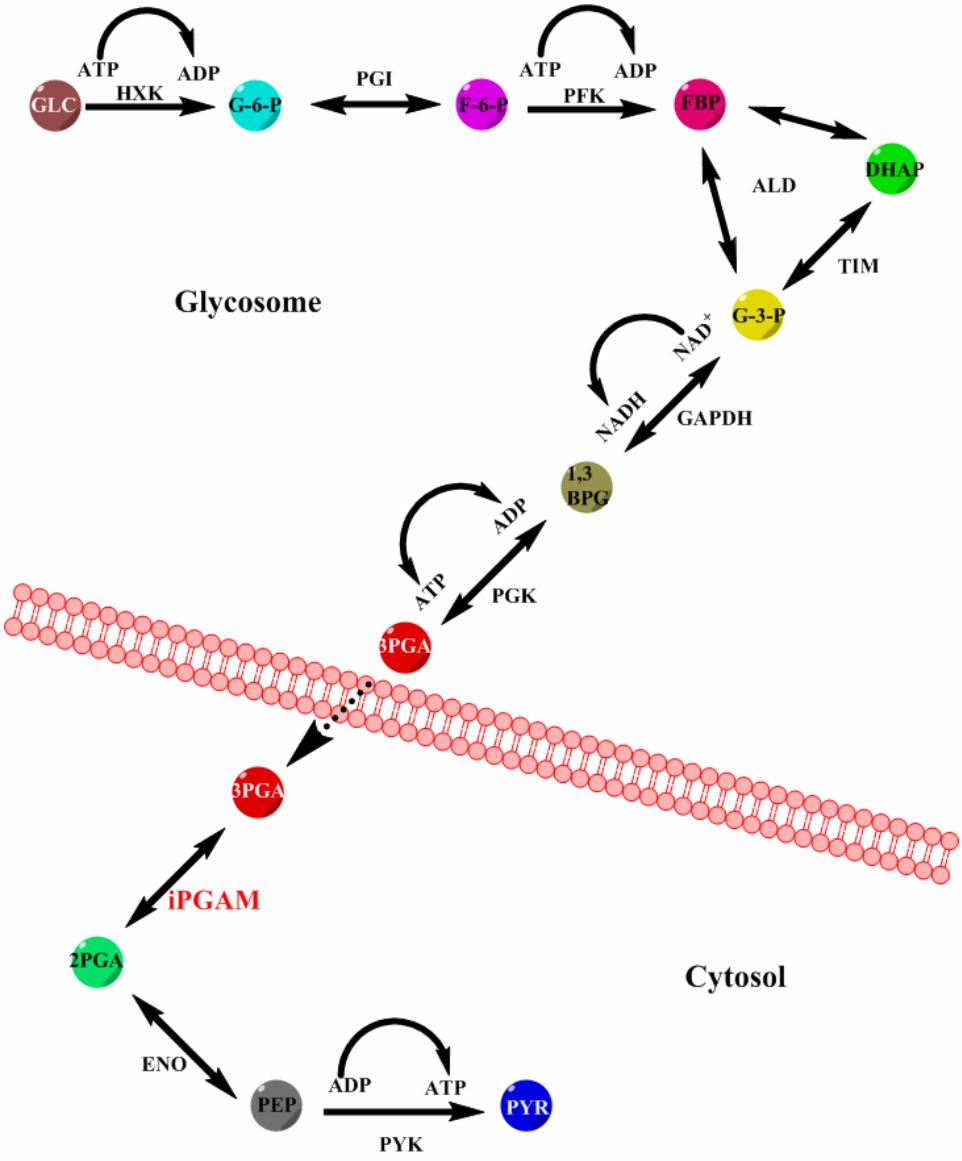


Figure 1.2 The glycolytic enzymes in trypanosomatids, which are located in the glycosome and in the cytosol.

Several enzymes in the glycosome, such as hexokinase (HXK) (Seed and Baquero 1965; Sharlow et al. 2010), PFK (Cronin and Tipton 1987; Michels et al. 1997; Martinez-Oyanedel et al. 2007; McNae et al. 2009), fructose 1,6-bisphosphate aldolase (EC 4.1.2.13) (Callens et al. 1991; de Walque et al. 1999; Dax et al. 2006) and PGK (Colasante et al. 2007) have been characterised both structurally and biochemically. This has shed further light on understanding their catalytic mechanisms, as well as their biochemistry, which are important steps in designing inhibitory approaches. Recent studies have also reported the discovery of novel compounds which may well be developed into leads for some of these enzymes, for example, HXK (Sharlow et al. 2010) and PGK (Bressi et al. 2000). The three final glycolytic enzymes in trypanosomatids which are located in the cytosol have also been characterised extensively. The iPGAM (Poonperm et al. 2003; Guerra et al. 2004; Nowicki et al. 2009; Fuad et al. 2011), ENO (Quiñones et al. 2007; Avilán et al. 2011) and PYK (Ernest et al. 1998; Rigden et al. 1999; Morgan et al. 2011) structural and biochemical studies have greatly progressed in recent years. Moreover the location of these enzymes may provide better access for inhibitors, rather than having to go through the glycosomal membrane to reach the target molecules.

It is indeed interesting to correlate drug design with the control of glycolytic flux. This has been shown in a study by Bakker et al. in 2000, where both the computer simulation and experimental studies exhibited the roles of several steps in glycolysis in controlling the glycolytic flux in *T. brucei*. Based on earlier studies by Bakker et al. 1997, a subsequent mathematical model was proposed by using experimentally-determined kinetics as the foundation (Albert et al. 2005). The outcome from this study revealed that the concentrations of five glycolytic enzymes (HXK, PFK, iPGAM, ENO and PYK) could be substantially decreased by RNAi experiments. Moreover, enzymes in the glycosome and the cytosol were all affected by the RNAi knockdowns, which further suggested a unique regulatory mechanism of glycolysis in trypanosomatid parasites. More recently, Achcar et al. 2012 revealed that two enzymes in particular, glyceraldehyde 3-phosphate dehydrogenase (GAPDH) (EC 1.2.1.12) and iPGAM

maximally controlled the glycolytic flux based on the control coefficients that have been analysed in the dynamic model. It is interesting too that the model in this study has been built by taking into account parameter uncertainty, which occurred due to experimental errors or lack of data, and thus increased the knowledge of control hierarchy that is important in prioritising potential drug targets (Achcar et al. 2012).

1.3 Phosphoglycerate mutase

Glycolysis occurs by a series of biochemical reactions catalysed by individual enzymes. There are ten separate reactions in the metabolic pathway, in which some are reversible, forming the basis of the gluconeogenic pathway which runs in the opposite direction. PGAM is the enzyme catalysing the seventh step of glycolysis, converting 3PGA to 2-phosphoglycerate (2PGA) (Figure 1.2). In humans and vertebrates, as well as yeasts and bacteria (Rigden et al. 2001; Chiba et al. 2012), the enzyme requires 2,3-bisphosphoglycerate as the cofactor for the catalytic function (dPGAM). By contrast, the iPGAM found in protists, including *T. brucei* (Chevalier et al. 2000) and *L. mexicana* (Guerra et al. 2004), as well as in plants (Grana et al. 1992; Wang et al. 1996; Zhao and Assman 2011) and some nematodes (Raverdy et al. 2007), bacteria (Jedrzejas et al. 2000a; Jedrzejas et al. 2000b), archaea (Johnsen and Schönheit 2007), filarial parasites *Dirofilaria immitis* (Li et al. 2011) and its *Wolbachia* endosymbiont and the same endosymbiont from *Brugia malayi* (Foster et al. 2009) are active in the absence of the cofactor (Guerra et al. 2004; Nowicki et al. 2009; Fuad et al. 2011). On the other hand, iPGAM requires divalent metal ions to support its activity. These two enzymes, although they essentially play the same role in the metabolic pathway, are not evolutionarily-related and do not share any sequence similarity (Djikeng et al. 2007).

i) Cofactor-independent phosphoglycerate mutase and metal requirements for enzymatic activity

The NCBI database lists 1232 iPGAM sequences (March 2012), across a fairly wide variety of organisms. Out of these, 16 sequences representing different bacterial, algal, nematode, fungal, kinetoplastid and plant species were aligned using the program CLUSTALW2 from the EMBL-EBI database (Figure 1.3). It can be seen directly that residues corresponding to important features of these enzymes are identical among the organisms, for example the sequences encoding the metal sites (green and yellow shading), as well as the phosphorylated Ser75 (pink shading) and other active-site residues (cyan shading). However, it is also a striking observation that there are regions which are conserved only among the plants and kinetoplastid iPGAMs, in particular the inserted Tyr201 (red shading) and hinge residues (gray shading). This observation indicates the possibility that iPGAMs could be divided into two different families, where trypanosomatid and plant iPGAMs are in one family, while bacterial, nematode, algal and fungal iPGAMs are in another family. The distinguishing properties appear to be related to the differences in catalytic mechanism involving metal ions, which will be discussed in chapter 4. Moreover, the similarity between plants and kinetoplastid iPGAMs is highlighted by the higher percentage of sequence identity between the two groups of organisms. Table 1.2 tabulates the overall percent sequence identity across the wide range of organisms in Figure 1.3, where the percentage of identities between plants and kinetoplastid iPGAMs are 53-58%. Meanwhile, the overall sequence identities between the bacterial, algal, nematode and fungal iPGAMs with kinetoplastid iPGAMs are found to be between 26 and 35%. Nevertheless, all iPGAMs from different organisms are reported to be monomeric enzymes of about ~60 kDa. A phylogenetic analysis for the cofactor independent PGAM is shown in Figure 1.4, which is based on the sequence alignment in Figure 1.3.

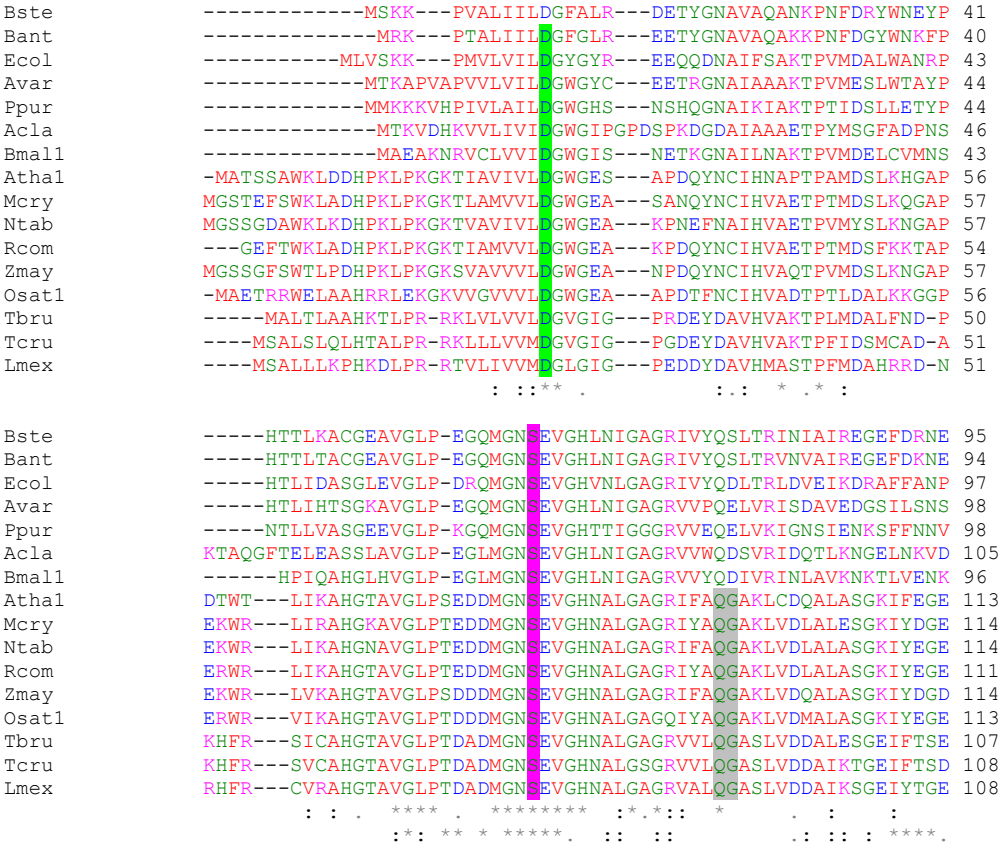


Figure 1.3 Alignment of iPGAM sequences, with residues coloured according to type. Abbreviations: Bste, *Bacillus stearothermophilus*; Bant, *Bacillus anthracis*; Ecol, *Escherichia coli*; Avar, *Anabaena variabilis* (blue-green alga) ; Ppur, *Porphyra purpurea* (red alga); Acla, *Aspergillus clavatus* (fungus); Bmal1, *Brugia malayi* isoform 1 (nematode); Atha1,*Arabadopsis thaliana* gene 1 (plant); Mcry, *Mesembryanthemum crystallinum*; Ntab, *Nicotiana tabacum*; Rcom, *Ricinus communis*; Zmay, *Zea mays*; Osat1, *Oryza sativa* subsp. japonica; Tbru, *Trypanosoma brucei*; Tcru, *Trypanosoma cruzi*; Lmex, *Leishmania mexicana*. Colour highlighting indicates the following: Metal site 1; Metal site 2; phosphorylated Ser75; other active-site residues; inserted Tyr201; hinge residues which are located in linker 1 and 2. The multiple sequence alignment was performed using CLUSTALW2 (http://www. ebi.ac.uk/ Tools/msa/clustalw2/) with the parameters set at default values. The analysis was performed by Dr. Linda Gilmore.

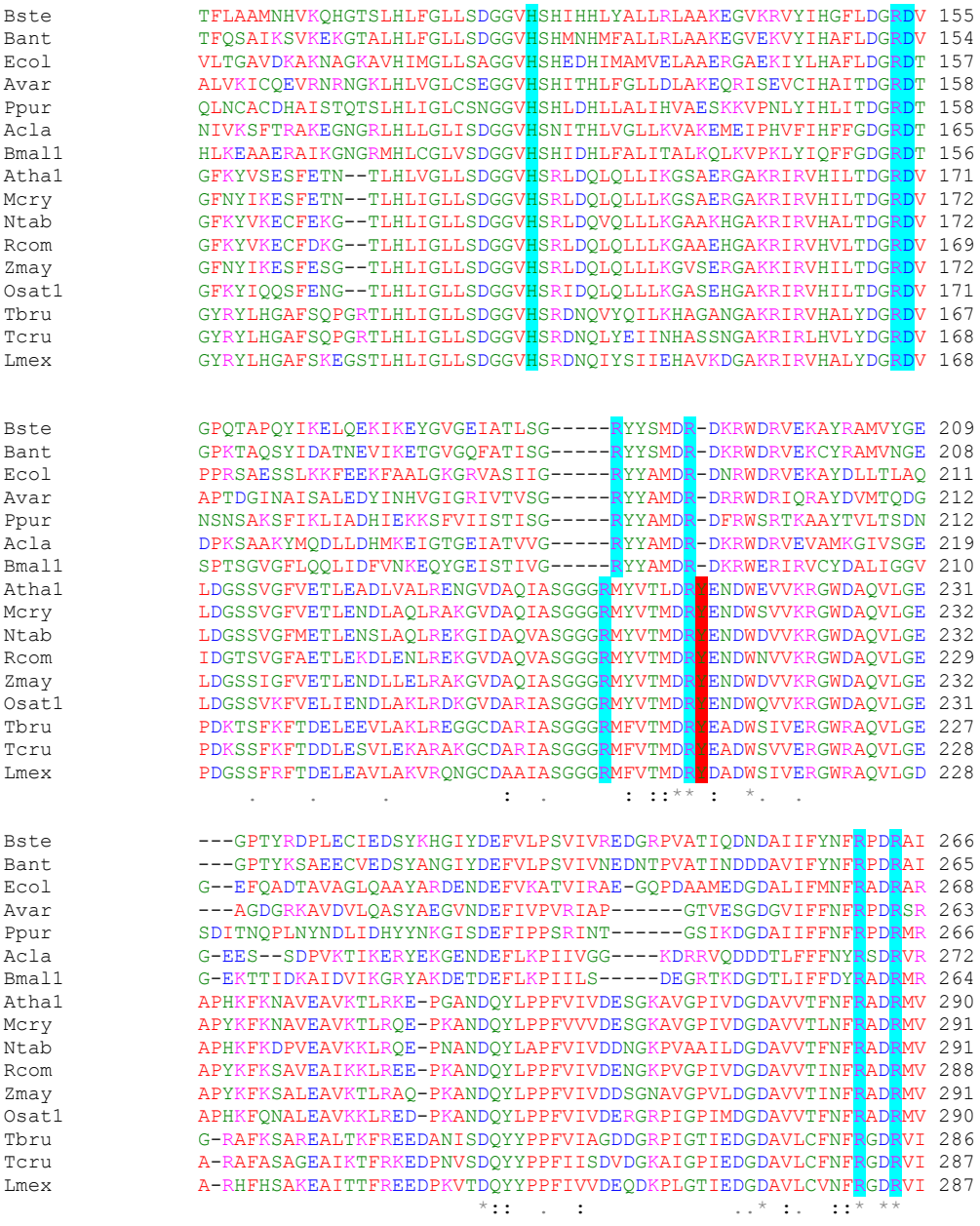


Figure 1.3 contd.

Figure 1.3 contd.

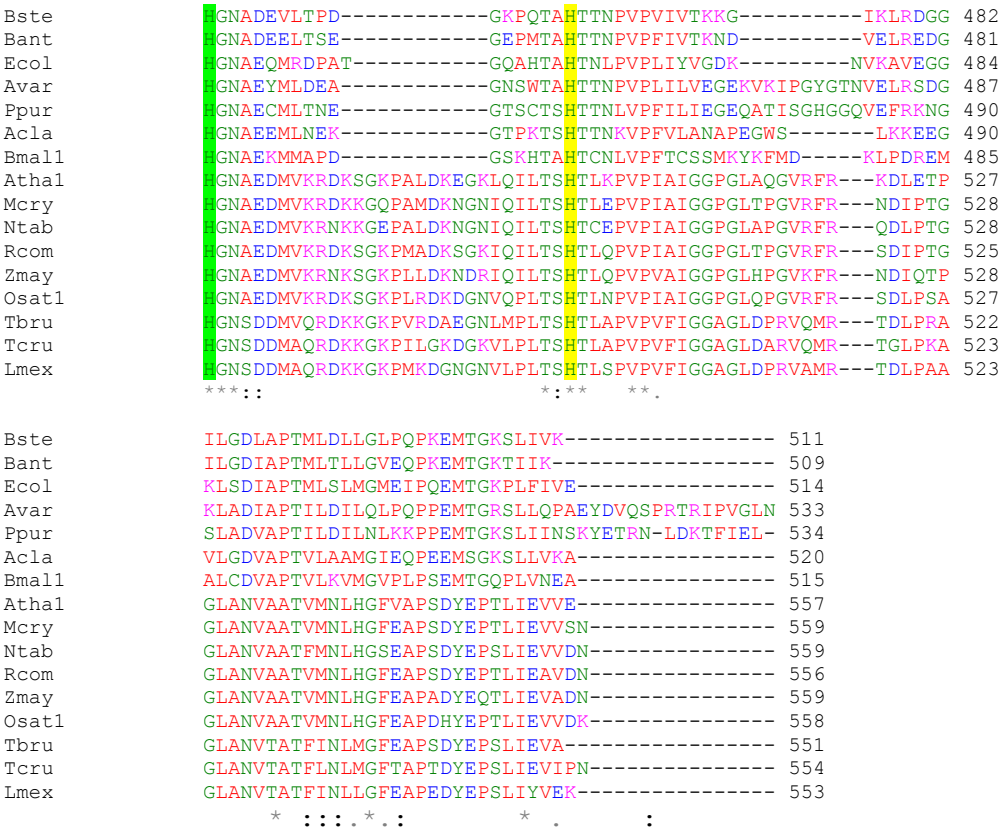


Figure 1.3 contd.

Table 1.2 Matrix of percent overall sequence identities among iPGAMs. Red shading indicates family 1 iPGAMs, with the darker shade corresponding to iPGAMs from bacteria. Blue shading shows family 2 iPGAMs, with the darker shade corresponding to iPGAMs from kinetoplastids. The analysis was performed by Dr. Linda Gilmore.

	Bste	Bant	Ecol	Avar	Ppur	Acla	Bmal	Atha	Mcry	Ntab	Rcom	Zmay	Osat	Tbru	Tcru	Lmex
Bste	100	77	48	48	45	45	40	33	33	34	33	32	33	35	33	33
Bant		100	49	47	42	43	40	32	31	33	31	31	31	35	33	33
Ecol			100	47	41	41	44	35	34	33	34	35	33	34	32	32
Avar				100	53	43	43	30	31	31	30	31	30	31	29	29
Ppur					100	39	40	29	28	26	27	27	27	29	29	27
Acla						100	53	30	30	30	30	31	30	29	30	29
Bmal							100	30	29	30	29	30	28	26	28	28
Atha								100	85	81	83	82	78	54	53	53
Mcry									100	84	88	85	80	56	55	54
Ntab										100	83	81	77	56	53	55
Rcom											100	82	81	56	54	54
Zmay												100	78	54	54	53
Osat													100	58	54	55
Tbru														100	83	77
Tcru															100	79
Lmex																100

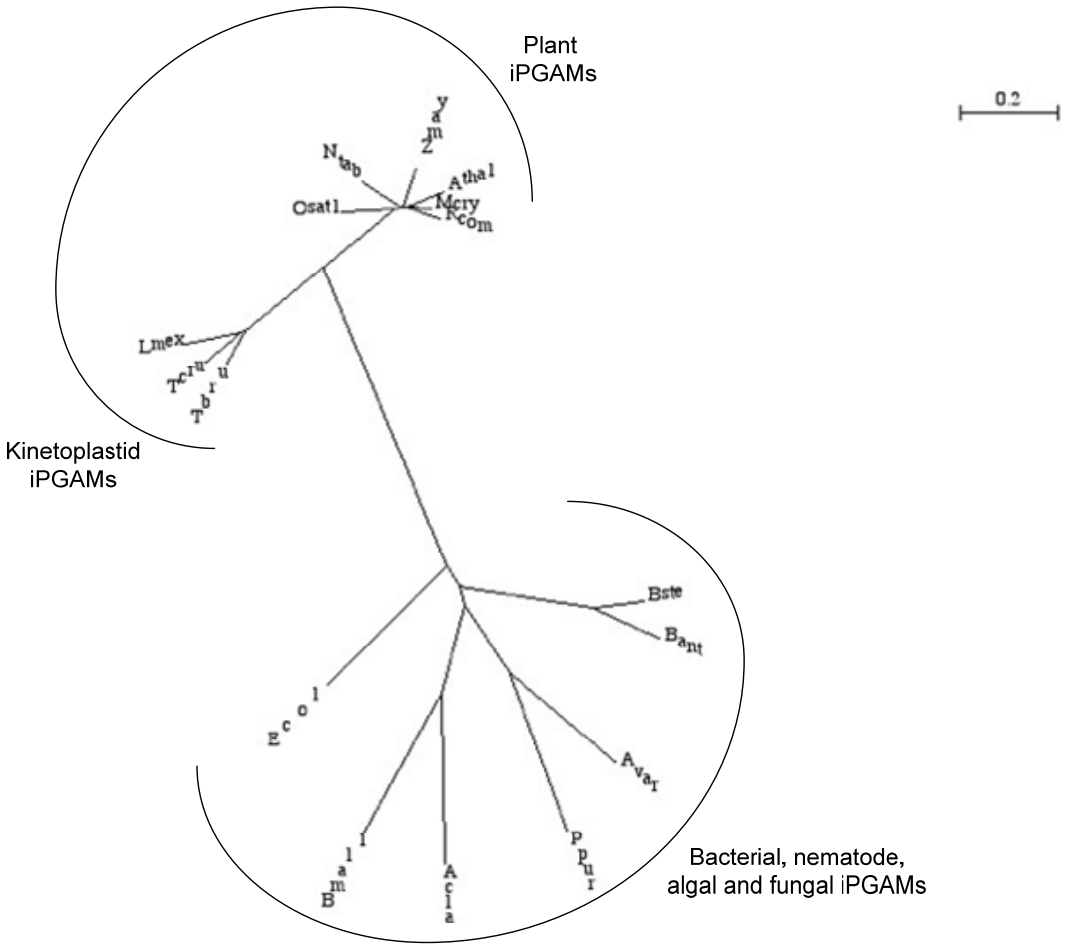


Figure 1.4 A phylogenetic tree of iPGAMs show the sequence relationships across a range of different species, and unite the plant and kinetoplastid iPGAMs under one family, while bacterial, fungal, nematode and algal iPGAMs are in another. The abbreviations for the species can be found in the legend for Figure 1.3. The analysis was performed by Prof. Fred Opperdoes.

The iPGAM metal requirement has been elucidated in various biochemical and structural studies (Guerra et al. 2004; Jedrzejas et al. 2000). This distinctive feature as well as structural comparisons categorised the enzyme under the metal-dependent alkaline phosphatase superfamily, uniting it with a few other enzymes, such as alkaline phosphatases (EC 3.1.3.1) (Anderson et al. 1975; Rej and Bretaudiere 1980) and sulfotases (EC 3.1.6) (Lassila and Herschlag 2008) with similar metal-binding sites and conserved structural folds (Galperin et al. 1998; Galperin and Jedrzejas 2001). The members of the family catalyse phospho transfer reactions, which involve basic biochemical processes including energy storage, biosynthesis, or replication of genetic material (Boyer et al. 1973; Vetter and Wittinghofer 1999; López-Canut et al. 2011). iPGAM from different organisms have been reported to require a particular divalent metal ion to support their catalytic functions (Jedrzejas et al. 2000; Collet et al. 2001; Guerra et al. 2004). Bacterial iPGAMs were shown to have a preference for Mn^{2+}, for example in *Bacillus* species (Jedrzejas et al. 2000a; Jedrzejas et al. 2000b; Nukui et al. 2007). *E.coli*, which interestingly has been shown to possess both iPGAM and dPGAM, was also reported to contain Mn^{2+} for the former enzyme (Fraser et al. 1999). The activity of iPGAM from nematodes: *Onchocerca volvulus, Brugia malayi and Caenorhabditis elegans* after EDTA treatment were reported to be effectively restored by Mn^{2+} and Mg^{2+} (Raverdy et al. 2007), while plant iPGAM, specifically *Triticum aestivum* (wheat) was reported to be reactivated by Mn^{2+} and Co^{2+} (Smith et al. 1986). Interestingly, iPGAM from kinetoplastids, *T. brucei* and *L. mexicana* were previously known to require Co^{2+} for their activities (Collet et al. 2001; Guerra et al. 2004), before it was revealed that Co^{2+} is essentially absent in the cytosol, where the enzyme is located (Fuad et al. 2011). The conserved metal-binding site residues among iPGAMs from different organisms fits beautifully with the discovery of two different families of iPGAM, which can be classified based on protein sequences, metal-mechanism and domain movements. In this case, kinetoplastid and plant iPGAMs possess a one-metal mechanism, which resulted in smaller open conformation, while the bacterial, algal,

nematode, and fungal iPGAMs have a larger open conformation, due to the two-metal mechanism used during catalysis.

The understanding of iPGAM metal requirements has improved in the last decade, especially as enzyme structural information has become available. In 2000, the crystal structure of *B. stearothermophilus* iPGAM in complex with 3PGA and two Mn^{2+} ions (Figure 1.5 a) was been solved at 1.9-Å resolution (Jedrzejas et al. 2000a), as was the complex with 2PGA with two Mn^{2+} ions (Jedrzejas et al. 2000b), solved at 1.7-Å resolution. These were the first crystal structures of iPGAM ever reported, thus shedding light on understanding the mechanism of action of iPGAM from a structural viewpoint. The two steps which are involved during catalysis: i) the phosphatase reaction which removes the phospho group from carbon 2 or 3, and vice versa while generating the phosphoserine intermediates in which both Mn^{2+} ions with distorted square pyramidal geometry play an important role, and ii) the phosphotransferase reaction that occurs as the phospho group is transferred from the enzyme back to the glycerate molecule. This mechanism is consistent with the existence of two distinct domains: the phosphatase and transferase. The residues involved during catalysis thus are located in the active site which involves residues in both domains. In 2007, the enzyme open-form structure in *B. anthracis* was solved (Nukui et al. 2007) (Figure 1.5 b) in the presence of a water molecule and two Mn^{2+} ions, showing structural differences from that of the closed-form enzyme from *B. stearothermohillus*. Nevertheless, both the individual domains and the catalytic mechanisms are similar, which is consistent with the high sequence identity observed between these two *Bacillus* species (77%) (Table 1.2). The *Ba*iPGAM open-form structure also highlighted a new coordination geometry of the two Mn^{2+} ions which differs from the previous *Bs*iPGAM structure, and thus allows a better understanding of pH-dependent regulation of iPGAM during spore formation and germination (Nukui et al. 2007).

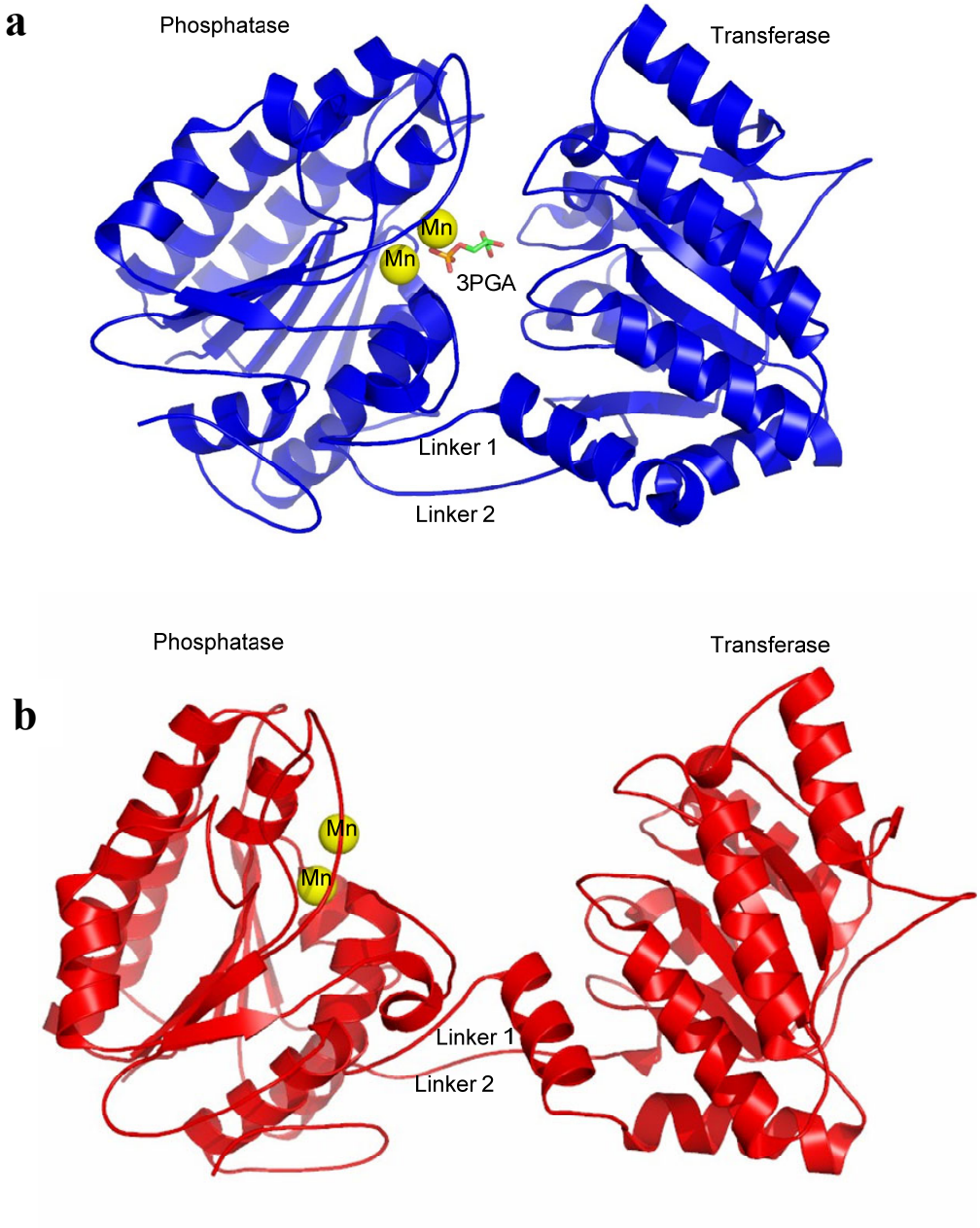


Figure 1.5 a)The closed-form crystal structure of *Bs*iPGAM in complex with 3PGA and two Mn^{2+} ions (Jedrzejas et al. 2000a, PDB ID 1EJJ) and b) The open-form crystal structure of *Ba*iPGAM in the presence of two Mn^{2+} ions and a water molecule (not shown) (Nukui et al. 2007, PDB ID 2IFY).

In recent years, our understanding of the structural aspects of trypanosomatid iPGAMs has significantly improved. The crystal structure of *Lm*iPGAM in complex with 3PGA and two Co^{2+} ions at high and low concentrations have successfully been solved at 2 and 1.9 Å resolution, respectively (Poonperm 2005; Nowicki et al. 2009) (Figure 1.6 a). The overall topologies are similar to the previously reported *Bacillus* iPGAM structures, where two domains are present which play individual roles for the catalytic function of the enzyme. The occupancies of both the substrate and product appeared to be equal in the active site, while both Co^{2+} ions are coordinated within the active site with different geometries and binding affinities. Co^{2+} in the first metal site (M1) possesses a distorted octahedral geometry with 100% occupancy, with a dissociation constant (K_d) value that is much lower than Co^{2+} in the second site (M2). Meanwhile, Co^{2+} in the M2 site possesses a tetrahedral geometry and coordinates with Ser75, a residue involved in phosphotransfer. The latter showed partial occupancy with an estimated K_d of > 1mM (Nowicki et al. 2009). One of the highlights from this structure is the observation that the side chain of Ser75 adopts an unfavourable position for catalysis when the M2 site is occupied, suggesting that catalysis occurs by a one-metal mechanism, instead of two as observed in the *Bacillus* iPGAM structure. Instead, Ser75 may be activated for phosphotransfer via H-bonding to the nearby residues. Another important structural feature which contributes to the difference from the bacterial enzyme is the existence of the inserted Tyr210 which causes the size of M2 to decrease, and correlates well with the lower metal-affinity. This was observed despite the conservation of 16 residues that are involved during catalysis (Figure 1.6b).

More recently, the open-form structure of *Tb*iPGAM has been solved at 2.3 Å resolutions in the presence of a sulphate and two Co^{2+} ions (Mercaldi et al. 2012). While the enzyme possess an α/β topology almost identical to *Lm*iPGAM with two individual phosphatase and transferase domains, the catalytic residues from both domains of *Tb*iPGAM lie farther apart, hence exposing them to solvent (see chapter 4). The structure is reported to represent the biologically relevant open conformation, and is strikingly less open than the previously reported *Ba*iPGAM open-form structure.

Nevertheless, this new structure also revealed that iPGAM possesses flexible hinges (Kempner 1993) that connect both domains, and thus confirms that the reaction mechanism is based on substrate-induced conformational change (Mercaldi et al. 2012). A highlight from the structure is the involvement of a metal-binding residue, Asp319, which coordinates the metal ion within the M1 site. The substitution of this residue with Ala resulted in complete loss of activity. It was thus concluded that the M1-metal/Asp319 interaction is important for the enzyme's structural function. Identical to previous structures, the 16 catalytic residues are conserved in the open-form *Tb*iPGAM.

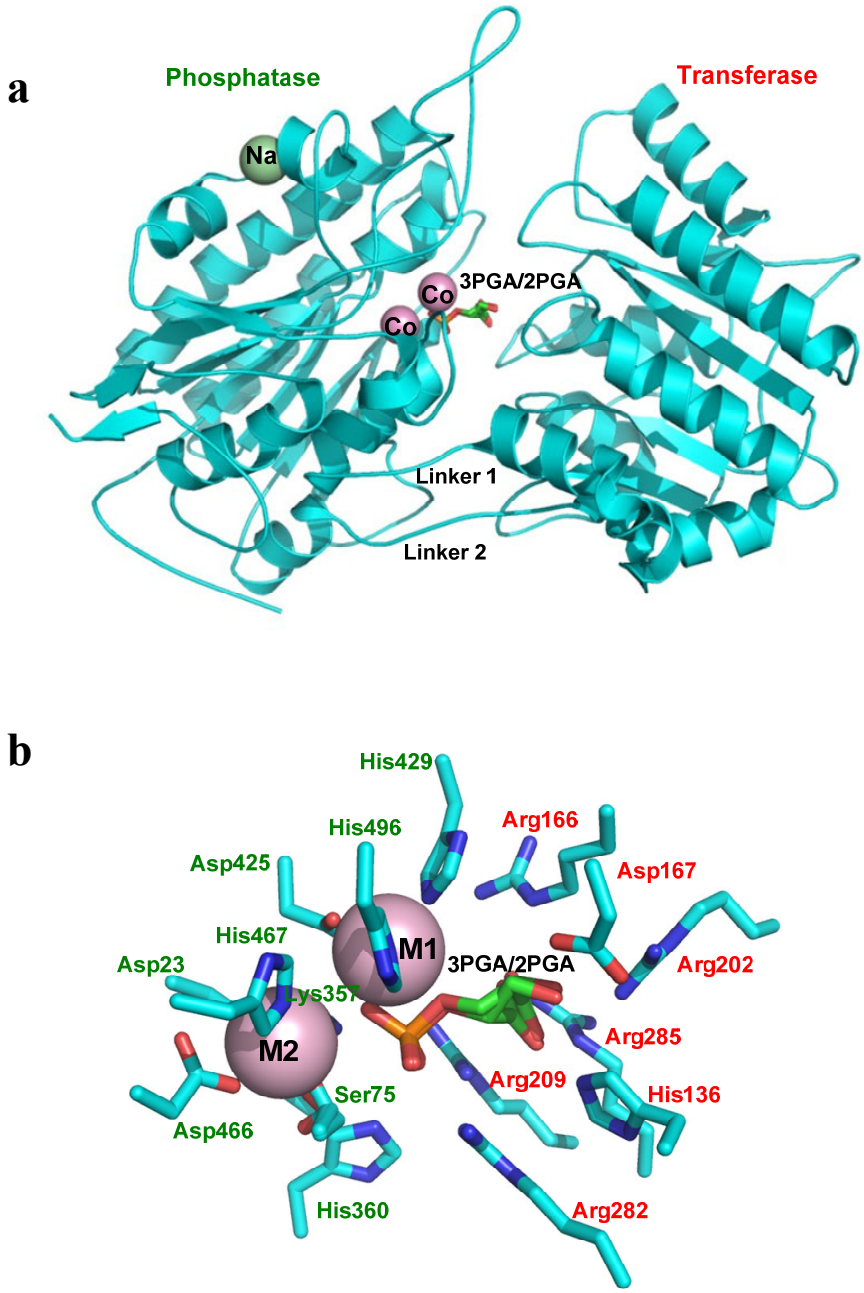


Figure 1.6 a) The closed-form crystal structure of *Lm*iPGAM (Nowicki et al. 2009, PDB ID 3IGY) in the presence of 3PGA/2PGA, two Co^{2+} ions and a Na^{+} ion which may be important for the enzyme structural integrity and b) 16 residues which are conserved in the active site. The green text indicates residues in the phosphatase domain, whereas the red text indicates residues in the transferase domain.

In an earlier section (see section 1.2), the potential of glycolytic enzymes as excellent drug targets has been discussed in general terms. Amongst all these enzymes, iPGAM may well be regarded as one of the key players in controlling glycolytic flux, as shown by the metabolic-associated mathematical model (Albert et al. 2005; Achcar et al. 2012). It is interesting to observe that iPGAM from *T. brucei* exhibits the lowest activity *in vivo* compared to the other enzymes, for example ENO (598 nmol. $min^{-1}.mg^{-1}$) and PYK (1020 nmol. $min^{-1}.mg^{-1}$) when measured experimentally (Albert et al. 2005), and this has also been shown in unpublished data obtained from the Brussels research group (Table 1.3). The low specific activity was also observed in this study, as discussed in chapter 5. The possibility of iPGAM as an important modulator for glycolytic flux makes it an exceptionally important drug target, and the characterisation of the enzyme, both from the structural and biochemical aspects is thus vital.

Table 1.3 The specific activities of trypanosomatid iPGAMs in the cytosol at *in vivo* conditions.

Enzyme	**Parasite stage**	**Specific activity ($nmol.min^{-1}.mg^{-1}$)**	**References**
*Tb*iPGAM	bloodstream	10-200	Chevalier et al. 2000
*Tb*iPGAM	bloodstream	50-80	Paul Michels (unpublished data)
*Tb*iPGAM	procyclic	486	Paul Michels (unpublished data)
*Tb*iPGAM	bloodstream	76	Albert et al. 2005
*Lm*iPGAM	promastigotes	300-400	Guerra et al. 2004
*Lm*iPGAM	promastigotes	672	Paul Michels (unpublished data)

ii) 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase

The 2,3-bisphosphoglycerate-dependent PGAM (dPGAM) is the corresponding evolutionarily-unrelated enzyme that can be found in humans and vertebrates, yeast, and several bacteria (Rigden et al. 2001; Chiba et al. 2012). Interestingly, some bacteria such as *E. coli* (Fraser et al. 1999), and archaea (Johnsen and Schönheit 2007) contain both the cofactor-dependent and independent PGAMs in their active forms, although the former has been reported to be the more predominant enzyme (Johnsen and Schönheit 2007; Rigden et al. 2001). While iPGAMs have been reported to be monomeric enzymes, dPGAMs are oligomers of 25 kDa subunits (Fothergill-Gilmore and Watson 1989; Rigden et al. 2001). As a more well-characterised enzyme, a significant number of crystal structures for dPGAM have been reported, for instance from yeast (Campbell et al. 1974; Crowhurst 1999) which was the first crystal structure reported for this enzyme, *E. coli* (Bond et al. 2001; Bond et al. 2002), Gram-positive pathogen *Mycobacterium tuberculosis* (Müller et al. 2005), human B type (Wang et al. 2005) and more recently from *Burkholderia pseudomallei* (Davies et al. 2011). The enzyme belongs to the dPGAM superfamily, together with acid phosphatases and fructose 2,6-bisphosphatases (Rigden et al. 2001). One of the most prominent features of dPGAM is the reversible inhibition by vanadate in the micromolar range (Carreras et al. 1980; Bond et al. 2002; Johnsen and Schönheit 2007). All dPGAM sequences contain a conserved histidine in the active site, which becomes phosphorylated during catalysis (Fothergill-Gilmore and Watson 1989; Rigden 2008). The lack of evolutionary relationship between trypanosomatid iPGAMs and human dPGAMs further augments the potential for the former enzyme to be an important drug target.

1.4 The biologically relevant metal ions

Metal ions are essential in many natural processes on earth, and play an important role in various biochemical processes in living cells. Most biochemical reactions such as cell signalling processes, enzyme catalysis and gene expression require metal ions for specific functions. Currently, there are at least 13 metals which are important for plants and animals (Bertini et al. 1994; Lippard and Berg 1994; Gomes and Wittung-Stafshede 2011). Out of this number, four of them: sodium (Na), potassium (K), magnesium (Mg) and calcium (Ca) are found to be abundant (Fenton 1995; Gomes and Wittung-Stafshede 2011), whereas the remaining nine are present in smaller quantities. These metals are known as trace metals, and consist of transition metal elements such as vanadium (V), chromium (Cr), molybdenum (Mo), manganese (Mn), iron (Fe), cobalt (Co), nickel (Ni), copper (Cu) and zinc (Zn). It is interesting that Fe^{2+} is the most widely-used trace metal, and for example reaches 4-5 g in human bodies (Fenton 1995; Gomes and Wittung-Stafshede 2011). Although the trace metals are present in cells and essentially important for various biological systems (Haraguchi 2004), unnecessary accumulation, analytical difficulties, the possibilities of contaminations at trace levels and the tendency of competition of different elements for the same chemical sites may contribute to the difficulties in understanding the nature of metals in a particular biological environment (Fraústo da Silva and Williams 2001). This thus affects the metal content and characterisation of metalloproteins, which depends on species, location, metal ion availability and redox states (Maret 2010). Hence it is important to analyse metal preferences by also bearing in mind the biologically-significant concentrations that may present in the biological system of interest.

i) The role of transition metal ions in biological systems

Homeostatic regulation of the transition metal ions is systematically controlled in living cells. This is vital to ensure that sufficient amounts of metals are present to assist metabolic functions, while at the same time preventing toxicity in the context of available binding sites in metalloproteins (Maret 2010). In biological systems, the order of trace divalent metal ion concentrations (excluding the bulk metals: Mg^{2+} and

Ca^{2+}) can generally be ordered as: Fe^{2+}, Zn^{2+} > Cu^{2+} > Mn^{2+} > Co^{2+}, Ni^{2+}. Table 1.4 tabulates the total metal concentrations measured in *E. coli*, humans, and the kinetoplastids *T. cruzi* and *L. major*. Each of the metals has distinct roles in explicit biological functions, and must be tightly buffered, in order not to interfere with the functions of other metal ions (Maret 2010).

Mg^{2+} is essential in all types of cells in organisms, and is reported to be the most abundant divalent metal ion that has been measured within cells (Andersson et al. 1997). Mg^{2+} is required for enzymatic reactions and structural roles (Cowan 1995), cell-signalling processes (Wu and Veillette 2011), nucleic acid biochemistry (Black and Cowan 1995) and also all reactions which require the hydrolysis of ATP. It is also found to be important in modulating cell proliferation, cell cycle progression and differentiation, and also is associated with the immune system (Tam et al. 2003). The wide functions of Mg^{2+} thus make it vital to be sufficiently available in all types of cells in animals, where the concentrations are found to be maintained at relatively different levels (Valberg et al. 1965; Iyengar et al. 1978).

Ca^{2+} plays an exceptionally important role in various biochemical processes, and has been well characterised, revealing its multiple functions in biological systems. For instance, it is important in controlling metabolic pathways, involved in common signalling mechanisms by binding to proteins such as calmodulin, apart from contributing to the mechanical stability of membranes, cell division and hormonal activities (Fraústo da Silva and Williams 1991; Lin et al. 2009). Zn^{2+} on the other hand, is involved in various catalytic and structural functions for enzymes, and plays a role in DNA regulatory functions in eukaryotic cells (Fraústo da Silva and Williams 1991). It is reported that Zn^{2+} can be found in almost 300 enzymes, which makes it significantly abundant in living cells, and widespread in nature (Lin et al. 2009).

Fe^{2+} is involved in various biological roles, for example in transportation of oxygen in haemoglobin, playing a role in electron-transfer and redox chemistry, or utilisation of hydrogen peroxide catalysed by the enzyme peroxidase (Fraústo da Silva and Williams 2001). Meanwhile, the role of Mn^{2+} has been elucidated mostly

by its wide functions as a cofactor in enzymatic reactions. One of the well known enzymes containing Mn^{2+} is superoxide dismutase (SOD) of prokaryotes and of the organelles of eukaryotes (Fraústo da Silva and Williams 2001). Earlier in this chapter, the role of Mn^{2+} as the cofactor for iPGAM has been discussed where it is vital for the enzyme's catalytic function (Jedrzejas et al. 2000a; Jedrzejas et al. 2000b). Cu^{2+} is best known for its involvement in electron transfer, as well as for assisting enzymatic functions such as transportation of oxygen, removal of hormones, or reduction of oxygen to hydrogen peroxides (Fraústo da Silva and Williams 2001).

Co^{2+} and Ni^{2+} are two of the transition metals that are found to be the least abundant in eukaryotic cells (Maret 2010). The functions of these two elements may represent a relic of early life before the advent of dioxygen (Fuad et al. 2011). The role of Ni^{2+} is very limited in higher organism, and is particularly well-known in urease. However, in free anaerobic bacteria such as methanogens, Ni^{2+} is important for the functions of some enzymes, for example Ni^{2+} dehydrogenases (Fraústo da Silva and Williams 2001). Meanwhile, Co^{2+} is important for the functions of vitamin B_{12} or cobalamin, which is necessary for various biological purposes, such as DNA synthesis, formation of red blood cells and maintenance of the nervous system (Karmi et al. 2011). Co^{2+} also plays a role as a component of the cofactor corrin (Harding et al. 2010). Nevertheless, Co^{2+} and Ni^{2+} are relatively rare in nature and may well be considered to be eliminated from biology, as neither seems to possess special chemistry in acid/base catalysis (Fraústo da Silva and Williams 2001), although they have recently been recognized to be essential in some living systems (Harding et al. 2010).

Table 1.4 Divalent metal concentrations in *E. coli*, *T. cruzi*, *L. major* and human.

Metal ions (total)	***E. coli* grown in LB medium (Outten and O'Halloran 2001; Finney and O'Halloran 2003; Maret 2010)**	***T. cruzi* epimastigotes as measured from acidocalcisomes (nmol/mg dry weight) (Docampo et al. 2005)**	***L. major* promastigotes as measured from acidocalcisomes (nmol/mg dry weight) (Docampo et al. 2005)**	**Cytosol of eukaryotes (Permyakov and Kretsinger 2009)**
Ca	10 μM-10 mM	171 ± 6	39 ± 18	1.5-1.75 mM
Co	Low abundance	n/a	n/a	Low abundance
Cu	10 μM	n/a	n/a	
Fe	100 μM	n/a	n/a	
Mg	> 10 mM	646 ± 19	515 ± 179	2.3 mM
Mn	10 μM	n/a	n/a	10-100 μM
Ni	Low abundance	n/a	n/a	Low abundance
Zn	0.1 mM	148 ± 6	74 ± 64	0.2 mM

ii) Metalloenzymes and the biochemistry of metals in protein structures

It has been reported that more than 30% of proteins in living cells utilise metals to perform their biological functions (Gray et al. 2003; Gomes and Wittung-Stafshede 2011), and 40% of all enzymes possess metals, either for structural or catalytic reasons (Bertini et al. 1994; Lippard and Berg 1994; Harding et al. 2010; Gomes and Wittung-Stafshede 2011). Most metalloenzymes have one or two metal binding sites that usually exist in the catalytic domains. The choice of metals which the protein prefers is predicted to depend upon the concentrations of the elements in the cells (Maret 2010), as well as the geometrical topology of the catalytic site that determines the interactions between the corresponding metal ions and the coordinating ligands (Harding et al. 2010). In proteins, atoms which usually interact with metal ions are imidazole nitrogen from histidine, sulphur from cysteine, carboxylate oxygen from aspartic and glutamic acids, and oxygen of threonine or serine, and also the main chain carbonyl oxygen from any amino acid residues. Different metal ions may have different ranges of K_d when it comes to binding affinity towards the ligands, which

may vary in orders of magnitude (Harding et al. 2010). The Irving-Williams series (Irving and Williams 1953) lists the interactions of metal ions and ligands in the following order of affinity: $Mn^{2+} < Fe^{2+} < Co^{2+} < Ni^{2+} < Cu^{2+} < Zn^{2+}$, where Mn^{2+} and Fe^{2+} were reported to bind rather weakly, while Zn^{2+} and Cu^{2+} bind more strongly (Maret 2010). Nevertheless, the affinity of the metals for proteins are considered to be high if the metals are required to be bound for specific functions (Maret 2010), which depend upon i) careful considerations on protein preferences and geometrical coordination of the metal atoms to the interacting protein and ii) the availability of the metals at a biologically relevant concentration.

Advances in protein crystallography have allowed detailed structural information to be revealed, but often the identification of specific metals in a structure requires knowledge of the preferred geometrical coordination. The 'first coordination sphere' refers to the group of atoms within chemical bonding distance of the metals (Harding et al. 2010). This determines the coordination number and the geometry that are preferred and commonly found for specific coordination groups, as tabulated in Table 1.5. However, substantial distortions from the idealised structures may occur too in metalloproteins (Bertini et al. 1994; Gomes and Wittung-Stafshede 2011). In some cases, there may well be more than one possibility when trying to identify a particular metal based on coordination with the neighbouring atoms. For example, Mg^{2+} and Co^{2+} both possess the coordination number 6, with octahedral as the ideal geometry. Interactions between the metals and the neighbouring atoms from the nearby residues may also depend on the possible metal-donor atom distances. In addition, the preference may well be associated with the ionic radii of the metals, in which specific metals may appropriately fit into a particular binding site. Once these possibilities are considered, correct assignments of metals may well be obtained.

Table 1.5 Metal preferences for coordination number, preferred coordination geometry and usual donor atoms for metal-protein interactions. This table was adapted from Harding et al. 2010.

Metals	Usual coordination number	Other coordination number	Preferred coordination geometry	Usual donor atoms	Other donor atoms
Ca^{2+}	6	4,5,7,8	octahedral	O of aspartate and glutamate O main chain	O of asparagine and glutamine O of serine and threonine, N of histidine
Co^{2+}	6	4,5	octahedral	O of aspartate and glutamate N of histidine	O of serine and threonine O main chain, S of cysteine
Cu^{2+}	3,4	5,6	square planar	N of cysteine S of cysteine	O of aspartate and glutamate O main chain, O of asparagine and glutamine
Fe^{2+}	5,6	4	octahedral/trigonal bypiramidal	N of histidine O of aspartate and glutamate	S of cysteine
Mg^{2+}	6	3,4,5	octahedral	O of aspartate and glutamate O main chain	O of serine and threonine O of asparagine and glutamine, N of histidine
Mn^{2+}	5,6	5,6	octahedral/trigonal bypiramidal	O of aspartate and glutamate N of histidine	O main chain, O of asparagine and glutamine O of serine and threonine
Ni^{2+}	6	4,5	square planar	N of histidine O of aspartate and glutamate	S of cysteine, O main chain
Zn^{2+}	4	3,5,6	tetrahedral	N of histidine S of cysteine O of aspartate and glutamate	O main chain, O of asparagine and glutamine O of serine and threonine

iii) Metals affecting protein folding and conformation

Crystallography is obviously an excellent method to obtain protein structural information. Over the years, numerous high-resolution metalloprotein structures have been solved, allowing better understanding of the occurring interactions at the macromolecular level. However, little is known about the biosynthesis of these proteins (Wittung-Stafshede 2002), or how living cells may be populated with proteins with weakly-bound metals, and at the same time, allow the existence of proteins that have tightly-bound metals (Gomes and Wittung-Stafshede 2011). Two main ideas have been proposed as explanations: i) the compartment (with its metal composition) where the protein folds may control the metal-binding preference and metal content (Tottey et al. 2008), and ii) metal insertion into proteins is controlled by specific or non-specific protein-chaperones or protein-based delivery systems (Gomes and Wittung-Stafshede 2011). Figure 1.7 illustrates the biosynthesis of metalloproteins and how it correlates with metal-binding.

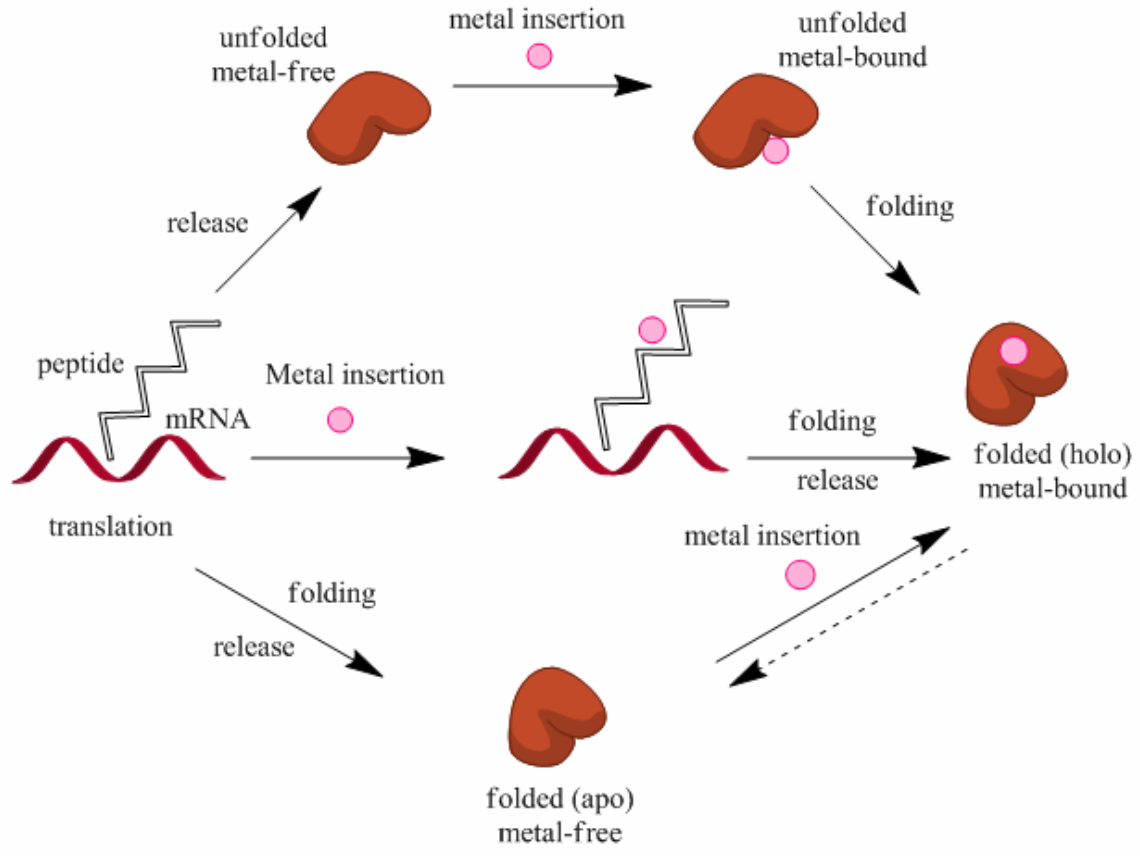


Figure 1.7 The illustration of metalloprotein biosynthesis and possible connections to metal binding. Metals could be inserted into the proteins at different stages of the protein biogenesis process. Metals insertion could occur either i) after the polypeptide has released from the ribosome, but before it is completely folded or, ii) after the polypeptide has folded into its conformation. The figure was adapted from Gomes and Wittung-Stafshede 2011.

It has been postulated that metals also play roles during protein folding and act as modulators for protein conformational changes. At *in vitro* conditions, it has been observed that metals are incorporated and bind into proteins after polypeptide unfolding (Robinson et al. 1997; Gomes and Wittung-Stafshede 2011), which thus suggested that *in vivo,* the metals may interact with proteins before the polypeptide folds. This may affect the folding reaction, in which protein misfolding may occur, due to specific coordination of the metal cofactor (Pozdnyakova et al. 2000). On the other hand, metals also play a role in structure stabilisation, and thus are essential for the protein's integrity, as observed in the Zn^{2+}-coordination in zinc-finger motifs

(Petsko and Ringe 2004). In some cases, metal-binding may also trigger aggregation to oligomerisation-prone proteins, which may cause oligomeric intermediates, non-native protein conformations and destabilisation of the molecules. This has been shown in some cases, for example, Cu^{2+} and Zn^{2+} ions have been observed to induce aggregation of the amylogenic peptide β-microglobulin (Villanueva et al. 2004; Blaho and Miranker 2009; Gomes and Wittung-Stafshede 2011). This phenomenon is prone to happen if the metals are present in large amounts or are freely available in biological fluids, under a non-physiological condition (Gomes and Wittung-Stafshede 2011).

Protein conformations may change too, due to binding of metal ions to a folded polypeptide. This affects the conformations of the proteins by creating cross-links between a specific set of residues in the protein structure, resulting in stabilisation and maintenance of a specific structure (Gomes and Wittung-Stafshede 2011). For example, human plasminogen activator inhibitor type 1 (PAI-1) was found to be stabilised by Ca^{2+}, Mg^{2+} and Mn^{2+}, but on the other hand, exhibited destabilisation in the presence of Co^{2+}, Cu^{2+} and Ni^{2+} (Thompson et al. 2011). Moreover, the stabilisation and destabilisation effects have been shown in S100A2 protein, a regulator of the cell-cycle, which was activated and stabilised by Ca^{2+}, and shown to be in the opposite condition by Zn^{2+} binding (Botelho et al. 2009). Conformational changes were also observed in CheY cytoplasmic protein, where the binding of divalent metal ions: Mg^{2+}, Ca^{2+}, Sr, Zn^{2+} and Mn^{2+} resulted in metal-free and metal-bound conformational changes (Kar et al. 1992). In another case, Zn^{2+} was shown to affect the conformational and oligomerisation pattern of the p53 tumor suppressor proteins (Coffer and Knowles 1994). Effects of different metal ions on the 3-dimensional structure of human carbonic anhydrase have also been reported in an earlier literature (Coleman 1965). These observations thus show the importance of metals in maintaining protein native structures for biological functions.

ii) Non-native metals affecting enzymatic activities

Almost one quarter of proteins require metal ions to maintain their structure and function (Harding et al. 2010), but many of the observations on the biologically

relevant metals that regulate enzyme activity have remained unclear. Metal content normally depends on the species, environment, metal availability and redox state, and these may affect the allocation of the correct metals in the protein structure. In some cases, the native metals that function *in vivo* may be substituted by non-native ones *in vitro.* This may also occur in organisms with few selectivity filters, which allows more diverse metal ions to enter the cell and bind to the binding sites, or when the native metals dissociate from the metal-protein complexes and are replaced by others (Maret 2010). Often too, metal-binding may be masked by metals that possess tighter binding affinity, and thus cause changes in the catalytic reactions (Maret 2010). Hence, the study of metalloproteins and enzymatic regulation by metal ions have always been a challenging task to researchers, in which sometimes there are no definite answers to the questions that were being raised, such as the proteins biologically relevant or native metal requirement .

Over the years, numerous studies have shown that there can be misleading information on the metal ions that are found in metalloproteins. In some cases, non-native metals can be incorporated into proteins during expression in the culture medium or in the purification procedure. The insertion of the artificial metals may substantially hyper activate or decrease the protein's activity, if the metals are present in the buffers and solutions. This phenomenon has been demonstrated in some cases, for example, the incorporation of Co^{2+} instead of Zn^{2+} into yeast alcohol dehydrogenase when the cells were grown in Co^{2+} (II) enriched medium (Curdel and Iwatsubo 1968). Metal substitution *in vitro* can cause hyper activation with the non-native metals, for example Co^{2+}, Cu^{2+} and Ni^{2+} insertion instead of Zn^{2+} in amino peptidase enhanced the activity of the enzyme by up to 10-fold (Prescott et al. 1985). By contrast, the substitution of Zn^{2+} by Co^{2+} in carbonic anhydrase in marine diatoms was reported to cause the enzyme to be less active than the native Zn^{2+} form (Yee and Morel 1996; Lane and Morel 2000), and is an example of *in vivo* metal substitution.

The correct assignment of metals in proteins that possess two metal sites is also of particular interest. Some proteins function with a one-metal mechanism for catalysis, but the presence of metals in the second site may affect the enzyme

function (Tainer et al. 1982; Holz et al. 2003). More explicitly, the regulation of enzymatic activity may vary, either by inhibitory or activating effects by the second metal, and may also determine substrate specificity (Sugiura and Noguchi 2009; Maret 2010). This has been shown in some cases, for example in carboxypeptidase A, where Zn^{2+} in the second metal site inhibits the enzyme as confirmed when the structure of the zinc-inhibited enzyme was solved (Gomez-Ortiz et al. 1997; Bukrinsky et al. 1998; Maret 2010). Another interesting example occurred when *in vitro* metal substitution provides illusory information about the metal requirement of an enzyme, such as in human phosphodiesterases. In this case, the enzyme appeared to be regulated by Mg^{2+} and Zn^{2+} with the latter metal exhibiting an inhibitory effect, but the interpretation was not correct because the former metal had been replaced by Mn^{2+} *in vitro* (Cowan 1997). These observations sound a note of caution in assigning the correct metals in proteins, or in distinguishing between two or one-metal mechanisms during catalysis (Maret 2010).

1.5 Overall aims of the project

This study concerns the molecular and structural properties of the enzyme phosphoglycerate mutase from *L. mexicana* and *T. brucei*, with the aim of obtaining substantial biochemical information in relation to drug design and discovery. For the purpose of obtaining structural and biochemical information that complement the previous *Lm*iPGAM closed-form structure, this enzyme has been the main focus in this project. The aspects of this study mainly cover:

1. Elucidation of the biologically relevant metal requirements of the trypanosomatid iPGAM under physiological conditions.

2. Biophysical characterisation with the aim of obtaining information on directing the *Lm*iPGAM towards the open-form conformation, which is the conformation desired for the purpose of drug design.

3. Drug design and inhibition studies to acquire knowledge for potential novel inhibitors, which ideally possess particular chemical properties that may resemble the substrate 3PGA or product 2PGA.

The chapters in this thesis cover various aspects to achieve these aims, as listed below:

1. Chapter 2 discusses the materials and methods that have been employed in this study.
2. Chapter 3 mainly revolves around the molecular biology, expression and new purification approaches for both *Lm*iPGAM and *Tb*iPGAM.
3. Chapter 4 covers various biophysical techniques in relation to distinguishing the different conformational and oligomeric states of trypanosomatid iPGAM in solution.
4. Chapter 5 elucidates the significance of metal content *in vitro* and *in vivo* in determining the biologically relevant metals for the trypanosomatid iPGAMs.
5. Chapter 6 focuses on the development of the multimode-plate reader discontinuous assay in providing biochemical information on the enzyme's metal requirements.
6. Chapter 7 covers the approaches utilised to find novel inhibitors for the trypanosomatid iPGAMs.
7. Chapter 8 concludes the main findings, discusses significant contributions and suggests future work.

Chapter 2:

Experimental

2.1 Experimental for Chapter 3

The pET28a plasmids encoding His-tagged *Lm*iPGAM and *Tb*iPGAM were obtained from Prof. Paul Michels, and all the steps outlined below are the same for both, unless otherwise stated.

2.1.1 Basic molecular cloning

The purpose of molecular cloning is to obtain the un-tagged version of the enzymes, hence all the steps outlined below is specifically designed for the un-tagged *Lm*iPGAM and *Tb*iPGAM.

i) Plasmid DNA isolation

A single colony from a freshly streaked selective plate containing pET28a plasmids with *Lm*iPGAM or *Tb*iPGAM genes was inoculated into 1-5 ml LB medium containing 50 µg/ml kanamycin (Melford Labs). The culture was incubated for 12-16 h at 37ºC with vigorous shaking before being harvested by centrifugation at 13,000 rpm in a table-top micro centrifuge for 3 min at room temperature. The plasmid DNA isolation was carried out using QIAGEN Plasmid Mini Kit (Qiagen).

ii) Digestion with restriction enzymes

Digestion with restriction enzymes was carried out in a 20 µl reaction mixture containing 1 µl 1X buffer (50 mM Tris-HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol), 3 µl plasmid DNA, 0.5 µl 100 mg/ml BSA, 1 µl 20,000 units/ml restriction enzymes and 14 µl water. The restriction enzymes that were utilised in this study were *Nco*I and *Xho*I for the construction of the un-tagged *Lm*iPGAM, and *Nde*I and *BamH*I for the construction of the un-tagged *Tb*iPGAM. The mixture was incubated at 37ºC for 2-4 h. A subsequent incubation was performed at 65ºC for 20

min to inactivate residual restriction enzymes that were left in the mixture. Agarose gel electrophoresis was carried out using 0.8% (w/v) agarose gel prior to ethidium bromide staining. The resulting digested plasmids were then analysed under the UV light.

iii) Gel extraction

The DNA fragments were excised from the agarose gel with a clean and sharp scalpel. The subsequent procedure was performed using QIAquick® Gel Extraction Kit (Qiagen).

iv) Ligation of DNA fragments

Ligation was carried out following gel extraction for both the vector and insert. For *Tb*iPGAM, the cloning process was carried out using Rapid DNA Ligation Kit (Roche). Table 2.1 tabulates the composition of the ligation mixture for the construction of the un-tagged *Tb*iPGAM constructs. The ligation mixture was incubated at 15ºC-25ºC for 5 min before being used for transformation.

Table 2.1 The composition of the ligation reaction, which consists of vector, insert, and the T4 DNA ligase enzyme.

Component	**Volume (µl)**
Vector (pET30a) (dilution in 5X dilution buffer *)	1.25
Insert (dilution in 5X dilution buffer *)	3-6
2X Ligation buffer **	1
U/µl T4 DNA ligase enzyme	1
Final volume	20

* DNA Dilution Buffer, 5x concentrated (Rapid DNA Ligation Kit (Roche))
** T4 DNA Ligation Buffer, 2x concentrated (Rapid DNA Ligation Kit (Roche))

Meanwhile, *Lm*iPGAM underwent a PCR cloning step prior to transformation, in order to clone the gene into the final vector. For this purpose, the CloneJET™ PCR Cloning Kit (Fermentas) was utilised to clone the amplified *Lm*iPGAM gene fragment into pJET 1.2/blunt cloning vector, which played a role as the intermediate vector. Two major steps were involved in the cloning process: i) the blunting reaction (Table 2.2) and ii) the ligation reaction (Table 2.3). The ligation mixture was incubated at room temperature for 5 min and was used directly for bacterial transformation.The *Lm*iPGAM gene was subsequently cloned into pET28a, as discussed in chapter 3.

Table 2.2 The composition of the blunting reaction of the PCR product (*Lm*iPGAM gene) before insertion into pJET 1.2/blunt vector.

Component	**Volume (µl)**
2X reaction buffer	10
PCR product	1-2
Nuclease-free water	17
DNA blunting enzyme	1
Total volume	18

Table 2.3 The composition of the ligation reaction for the blunt PCR product (*Lm*iPGAM gene), for insertion into pJET 1.2/blunt vector.

Component	**Volume (µl)**
Blunting reaction mixture	18
50ng/µl pJET 1.2/blunt cloning vector	1
5 U/µl T4 DNA ligase	1
Final volume	20

v) Primers for PCR

In the process of constructing the un-tagged versions of both *Lm*iPGAM and *Tb*iPGAM, a set of primers for gene amplification was designed. PCR amplification for *Lm*iPGAM was essential for cloning, but for *Tb*iPGAM the PCR was used analytically. Both the forward and reverse primers (Table 2.4) were constructed based on *Lm*iPGAM and *Tb*iPGAM gene sequences and have been designed by the program Primer3.

Table 2.4 The set of primers which were used to amplify *Lm*iPGAM and *Tb*iPGAM genes.

Primer sequences	**Size of the amplified product**
*Lm*iPGAM	
F-5' TAGACTCCATGGATGGCAGCACTTCTCTTGAAGC-3'	1.6 kb
R-5' GTTGCACTCGAGTCACTTCTCGACGTAGATCAGGCT-3'	
*Tb*iPGAM	
F-5' ATGCGCTACGATGGTGACTTGGTATTC-3'	0.7 kb
R-5' AATGAGCGAACGGTTCGTAGTCTGATGGA-3'	

vi) Transformation

Competent cells were prepared using the rubidium chloride method (Promega). Transformation was carried out with the addition of 20 ng of recombinant plasmid into a vial of competent *E. coli* cells strain dH5α (for cloning) or various cell lines (for protein expression, see next section). The mixture was subsequently incubated on ice for 5 min before been heat-shocked for 45 s at 42ºC in the heating block. The tubes were immediately transferred onto ice just before the addition of 250 µl of SOC medium before being incubated at 37ºC with slow agitation for approximately 1 h. The cells were finally spread onto pre-warmed selective LB plates containing 50 µg/ml kanamycin and incubated overnight at 37ºC. As a negative control where

possibilities of contaminants could be observed, an aliquot of cells was included through the whole transformation process in the absence of added plasmids.

2.1.2 Initial culture and protein expression

A single colony from an overnight-grown LB plate was inoculated into 50 ml LB broth containing 50µg/ml kanamycin. The inoculated culture was left overnight in the incubator shaker at 250 rpm. Following the overnight incubation, 1/50 from the overnight culture was re-inoculated into fresh LB medium containing 50µg/ml kanamycin and left to shake again at 250 rpm. Once the OD_{600} reached 0.5-0.7, the expression was induced by adding a final concentration of 1 mM of IPTG (Melford Labs). The expression system for *Lm*iPGAM has already been established in the Edinburgh Structural Biochemistry Group, and this study has utilised the same conditions as follows: BL21 (DE3) used as the cells line, with incubation at 17°C for 20 h (Poonperm 2005).

Meanwhile, *Tb*iPGAM expression trials were carried out with four different cell lines. Table 2.5 indicates the cell lines used in the expression trials and the antibiotics required for each cell line. A volume of 1-2 µl (30-40 ng/µl) of His-tagged and un-tagged *Tb*iPGAM plasmids were transformed into the cell lines, and the colonies from the plates were inoculated overnight in 50 ml Falcon tubes containing a volume of 5 ml LB or 2XYT medium before being left to shake at 37°C at 170 rpm. Final concentrations of 1 mM IPTG were added once the OD_{600} reached up to 0.5 to 0.7. The cultures were subsequently left at three different temperatures: 37°C, 30°C and 18°C. Samples in the absence of IPTG were included as negative control.

Protein expression analysis was carried out by taking a 1 ml samples which had been cultured at the different temperatures. At 37°C, an aliquot of the culture was taken every hour until 6 h. For both 18°C and 30°C, the same procedure was conducted, but only for every 3 h, 6 h, 14 h and 23 h. The cultures were centrifuged at 13,000 rpm for 10 min in order to analyse the levels of expression. At this stage,

PBS buffer was added to resuspend the cells prior to cell lysis. Both enzymatic and mechanical (sonication) approaches were carried out in the cell lysis procedure.

Table 2.5 The cell lines and antibiotics used in *Tbi*PGAM expression trials.

Cell lines	**Antibiotics**
BL21 (DE3)	Kanamycin
Rosetta 2 (DE3) pLysS	Kanamycin, Chloramphenicol (Duchefa Biochemie)
BL21 (DE3) Gold	Kanamycin, Tetracyclin (Melford Labs)
BL21 (DE3) CodonPlus	Kanamycin, Tetracyclin, Chloramphenicol

In the enzymatic method, a final concentration of 4 mM $MgCl_2$ was added into the lysate, followed by 1 mg/ml lysozyme (Fluka Biochemica). 0.1 mg/ml DNase1 from bovine pancreas (Sigma-Aldrich) was subsequently added before the tubes were inverted gently and left in the rotator at room temperature (approximately 20°C) for 2 h. The samples were sonicated twice, with 10 s bursts and 20 s rest. The lysate was centrifuged for 10 min, 4 °C at 13, 000 rpm to separate the soluble and insoluble fractions. The supernatant and pellet, corresponding to the soluble and insoluble proteins were analysed by 12% SDS-PAGE (see section 2.1.4 ii).

2.1.3 Protein purification

i) Cell lysis

The cell pellet from a litre of culture was resuspended with Buffer A (50 mM NaH_2PO_4 (AnalaR Normapur®), 300 mM NaCl (Fisher Scientific) pH 8.0) in the presence of two tablets EDTA free protease inhibitors (Roche), until it reached a final volume of 25 ml per 5 g of cell pellet. After the pellet was fully dissolved in the buffer, the cells were lysed in a Constant System TS Cell Disruptor set to a maximum pressure of 25 000 psi. The lysate was subsequently centrifuged at 4°C, 23,000 rpm for 45 min, using a JA 25.50 rotor. The supernatant was filtered through a 0.22 μm filter before being loaded onto a 5 ml IMAC Fast Flow^{TM} (GE Healthcare

Life Sciences) column that had been pre-charged with 0.1 M Ni^{2+} (GE Healthcare Life Sciences) on an ÅKTAPurifier™ instrument (GE Healthcare Life Sciences), which was controlled by Unicorn™ version 5.1.1 software. All purification steps were performed using the same liquid chromatography instrument with the temperature maintained at 10°C. It is noteworthy that only the His-tagged enzymes were purified in this study, and hence discussed in the following sections.

ii) First step: Immobilised metal affinity chromatography

As the first step of purification, Immobilised Metal Affinity Chromatography (IMAC) was performed by exploiting the N- and C-terminal His-tags that are present in *Tb*iPGAM and *Lm*iPGAM, respectively. This step is a classical primary chromatography method, and performed on the basis of manipulating the affinity of His-tags towards divalent metal ions which are immobilised in the column. Prior to the purification step, the un-charged column was washed with three column volumes (CV) of water before being charged with 5 CV of 0.1 M $NiSO_4$ (Sigma Life Sciences). Subsequently, the column was thoroughly washed with 3 CV of water, before the equilibration step. If multiple runs have been carried out using the column, thorough cleaning must be carried out by stripping the metals with EDTA-containing buffer (0.1 M EDTA (Fisher Scientific), 0.5 M NaCl and 20 mM Tris-HCl (Invitrogen) pH 7.9) for 5 CV, followed by subsequent cleaning with water and buffer for 10 CV each. The column would then be recharged with the chosen metal ions for 5 CV. It is also important to prepare the column for sample application by equilibrating it with buffer A (see below) for 5 CV. All steps were performed by maintaining the flow-rate at 3 ml/min during the whole run. Following sample injection, the non-binding molecules were washed with buffer A for 20 CV (50 mM NaH_2PO_4 buffer pH8, 300 mM NaCl and 20 mM imidazole (Fisher Scientific) and recognised as the flow-through, followed by a long wash for 20 CV with 5% buffer B to wash out non-specifically binding molecules. *Lm*iPGAM was then eluted in a step-wise fashion, by increasing the imidazole concentration with buffer B to 500 mM. The corresponding peak fractions from the elution profile (see chapter 3) were pooled before being applied to the desalting column.

iii) Second step: Desalting

In order to remove unwanted smaller molecules, such as imidazole and excessive salt in the protein samples, a desalting step was employed using a 53 ml HiPrep 26/10 desalting column (GE Healthcare Life Sciences). This approach is a convenient method in a typical buffer-exchange procedure, which is also an alternative to the dialysis method. Prior to sample application, the column was equilibrated with desalting buffer containing 20 mM HEPES (Sigma Life Sciences) pH 7.6, for 2 CV. The flow-rate was 3 ml/min through the whole run. Sample application was performed by applying 8 ml samples on to the column. The eluted fractions were subsequently measured for protein concentration before being applied onto the anion-exchange column.

iv) Third step: Ion-exchange chromatography

Charge distribution differences on the surface of protein molecules enable proteins to be separated using ion-exchange chromatography. In this experiment, a 1 ml MonoQ 5/50 GL (GE Healthcare Life Sciences) anion-exchange column was utilised. In this particular step, the interaction between the negatively-charged residues on the surface of the proteins and the positively-charged resin resulted in the binding of the protein of interest. The high maximum back pressure for the column would only allow a low flow-rate to be employed during running, which was maintained at 0.5 ml/min. Prior to sample application, the column was equilibrated with 5 CV of washing buffer (buffer A) containing 20 mM HEPES pH 7.6. The maximum capacity of proteins that could be loaded is 50 mg for each run. Following sample application, the washing step took place by using buffer A for 2 CV before the elution step with buffer B (20 mM HEPES, 1 M NaCl pH 7.6) took place in a linear gradient for 80 CV. The resulting separated peaks eluted at this stage were pooled separately and buffer-exchanged using a desalting or PD10 column into a storage buffer containing 20 mM TEA-HCl pH 7.6 and 50 mM NaCl.

2.1.4 Protein determination

i) Protein concentration and UV absorbance at A_{280}

Pure proteins were concentrated by using a Vivaspin concentrator with a 10 000 molecular weight (MW) cut off by centrifugation at 13, 000 rpm at 4°C, to achieve the highest concentration of proteins. Subsequently, the concentrations of purified *Lm*iPGAM were determined by UV absorbance at A_{280} nm using a Jasco V-550 UV-VIS spectrophotometer determined by the standard Beer-Lambert law, with the known extinction coefficient (42 080 M^{-1} cm^{-1}) and molecular mass (61788.5 Da), which had been calculated by the Expasy ProtParam software (http://web.expasy.org/cgi-bin/protparam/protparam). The calculation of the protein extinction coefficient by the Expasy ProtParam software was performed based on the amino acid sequences following the protocol from Gill and von Hippel 1989. The protein was stored in the cold room (4°C) for further biochemical and biophysical assays or at -20°C with the presence of 10% glycerol.

ii) SDS-PAGE analysis

A standard procedure for examining the purity of the proteins was performed using SDS-PAGE analysis. This technique allows proteins to migrate through polyacrylamide gel pores after reaction with SDS. For iPGAM, 12% SDS-PAGE was chosen as the optimal gel percentage to observe significant separation of the protein bands. The following chemical components were used to prepare the 12% SDS PAGE gel; 30% acrylamide (Severn Biotech Ltd), 10% sodium dodecyl sulphate (SDS) (Sigma Life Sciences), 10% ammonium persulphate (APS) (Sigma Life Sciences) and TEMED (Sigma Life Sciences). Both the resolving and stacking buffers contained Tris-base (Invitrogen) with pH adjusted to 6.8 and 8.8, respectively, while the running buffer contained 250 mM Tris, 2 M glycine (Sigma Life Sciences) and 10% SDS pH 8.3. Samples were prepared by mixing 5 µl 2-4 µg proteins with 5 µl of 2X loading buffer (Tris-base buffer pH 6.8, 10% SDS, 10% glycerol (BDH Prolabo) and bromophenol blue (Sigma Life Sciences) with the addition of 1mM of freshly prepared DTT (Melford Labs). The mixture was then

heated at 100°C for 5 min before being loaded into the wells. The electrophoresis took place at room temperature at 200V for 1 h and 5 min.

2.2 Experimental for Chapter 4

2.2.1 Computational programs

PyMOL, MODELLER, HYDROPRO and STP

The program PyMOL (DeLano 2002) was used to generate protein structure diagrams. The open-form model of *Lm*iPGAM was produced using MODELLER from Friend (Abyzov et al. 2005), by utilising the open-form *Ba*iPGAM as the available model structure. The predicted hydrodynamic radii of the protein molecules were obtained from the program HYDROPRO (Garcia de la Torre et al. 2000; Ortega and Garcia de la Torre 2007; Ortega et al. 2011) by utilising the molecular masses of *Lm*iPGAM, *Tb*iPGAM and *Ba*iPGAM separately and the specific volume of water (0.974 cm^3/g). The rest of the parameters were set up as default values. The program Surface Triplet Propensities (STP) (Mehio et al. 2010) was used to predict the most likely areas on the protein surface that may interact with another protein molecule for them to form a dimer. The analysis was performed by uploading the ligand-free *Lm*iPGAM PDB file into the program, which automatically coloured the the most or least likely areas on the protein surface that may interact with another macromolecule. The program can be accessed through the CTCB website with the link: http://opus.bch.ed.ac.uk/stp/.

2.2.2 Continuous coupling enzyme assay

iPGAM activity was assayed by monitoring the conversion of 3PGA to 2PGA, coupled with the oxidation of NADH to NAD^+ by LDH via ENO and PYK. Assays done during the purification of *Lm*iPGAM used 10 µl samples containing 0.2 mg/ml iPGAM which were added into a cuvette to give a final reaction mixture of 1000 µl containing 100 mM TEA-HCl buffer pH 7.6, 1.5 mM 3PGA (Sigma-Aldrich), 5 mM $MgCl_2$ (Fisher Scientific), 50 mM KCl (Fisher Scientific), 0.2 mM NADH (Roche), 1 mM ADP (Sigma-Aldrich), 2 units of ENO from bakers yeast (Sigma-Aldrich), 4

units of PYK from rabbit muscle (Sigma-Aldrich) and 6 units of LDH from rabbit muscle (Fluka Analytical). Controls were carried out with the addition of 2PGA into the reaction mixture, in the absence of *Lm*iPGAM and 3PGA. The decrease in A_{340} nm was monitored using a Jasco V-550 UV-VIS spectrophotometer and was used to obtain the rate of reaction for specific activity measurements (one unit corresponds to the conversion of 1 µmol of substrate min^{-1} mg^{-1} protein under standard conditions) (Fuad et al. 2011).

2.2.3 Gel filtration

Gel filtration is a conventional macromolecular separation technique that is performed on the basis of protein molecular masses and 3-dimensional shapes. In this study, analytical gel filtration was employed to characterise the distinct peaks eluted from the ion-exchange column. 5-10 mg of concentrated *Lm*iPGAM in 200 µl were applied onto a 24 ml Superdex 200 10/300 GL (GE Healthcare Life Sciences) column, which had been equilibrated for 2 CV with 20 mM TEA-HCl pH 7.6 and 50 mM NaCl. The column has the ability to separate molecules in a range of 10 kDa to 600 000 kDa, thus iPGAM which is ~60 kDa in size is a suitable candidate. The flow-rate was maintained at 0.5 ml/min, thus resulting in a total run time of ~50 min. The estimated molecular weights of the eluted fractions were based on the previous calibration of the column by a range of protein standards with varying molecular weights (see Appendix I (c)).

2.2.4 Dynamic Light Scattering (DLS)

In this study, DLS experiments were employed by analysing 50 µl of 1 mg/ml iPGAM that had been centrifuged for 10 min at 13, 000 rpm, and filtered through a 0.2µm filter. The samples were analysed using a 384 well plate, and the temperature for the analysis was set to be 10°C. The buffer containing 20 mM TEA pH 7.6 and 50 mM NaCl has to be freshly prepared. This analysis was performed using a Malvern instrument and analysed using the Malvern Zetasizer® software version 6.20. The data were expressed as the mode values of hydrodynamic diameter of the protein molecules based on intensity or volume distribution.

2.2.5 Native-PAGE

Native-PAGE analysis allows proteins to be separated based on the proteins shapes and conformational states, as well as the charge distribution on the surface of the macromolecules. In this experiment, 9% polyacrylamide gel was prepared using the chemicals described previously in section 2.1.4 ii, but in the absence of SDS. The running buffer contains 250 mM Tris and 2 M glycine pH 8.3, which allows the forward movement of iPGAM with a lower pI of 5.62. Samples were prepared by mixing 5 µl of 5-10 µg proteins with 5 µl of 2X loading buffer (see section 2.1.4 ii, without the presence of SDS and bromophenol blue) with the addition of 1mM of freshly prepared DTT. The electrophoresis took place in the cold room at 4°C, and set up to be running at 100V for 2 h and 10 min.

2.2.6 Size-exclusion chromatography coupled with Multi-Angle Light Scattering (SEC-MALS)

SEC-MALS analysis was performed to characterise iPGAM on the basis of its absolute molecular mass. This technique offers an advantage by combining macromolecular separation by size-exclusion chromatography, with the absolute molecular mass values from the MALS detector. Two other detectors are also important to quantify the concentration of the macromolecules, which are the Refractive Index detector (RI) and the UV detector. In this experiment, 40 µg of iPGAM was applied onto Superdex 200 10/300 GL column following the experimental conditions that had been established for the previous gel filtration analysis (see section 2.2.3). The Ettan™ LC (GE Healthcare) instrument was connected to a DAWN HELIOS II™ multiangle light scattering (MALS) instrument with 18 detectors and an Optilab® T-rEX refractometer (Wyatt Technology Corp., Santa Barbara, CA), which was controlled by the ASTRA light scattering software version 5.3.4 (Wyatt Technology Corp., Santa Barbara, CA). The analysis was performed at room temperature, approximately at 21°C. The data processing followed an earlier method as described in Folta-Stogniew and Williams 1999. The results obtained were expressed as the absolute molecular mass of the protein molecules in solution. This analysis was performed by Dr. Martin Wear and Dr. Elizabeth Blackburn.

2.2.7 MALDI-TOF mass spectrometry and ICP-MS

Mass spectrometry analysis, specifically MALDI-TOF is an important tool in characterising the masses of bio molecules, which are known to be easily fragmented with the traditional ionisation methods, as well as arbitrating the elemental composition in biological samples. Meanwhile, ICP-MS is also a type of mass spectrometry which is greatly sensitive, and possess the ability to determine the concentrations of metal ions at a range below one part per trillion. Both analyses are based on the ionisation of particles that are present in the sample, thus resulting in the generation of charged molecules. Molecules (or metals for ICP-MS) with different mass-to-charge ratios were then separated and measured by a mass spectrometer (Sparkman 2000). In the case of ICP-MS, the signal was proportional to the concentrations of the metals. The experimental protocols for both analyses are described below.

i) MALDI-TOF mass spectrometry

The MALDI-TOF analysis was performed by mixing a 0.5 µl aliquot of 1 mg/ml iPGAM with 0.5 µl sinapinic acid matrix, before loading onto a Voyager DE-STR MALDI-TOF gold target plate. The plate containing the samples was placed in a Voyager DE-STR MALDI–TOF mass spectrometer (Applied Biosystems, Warrington, Cheshire, U.K.). The mass map for the proteins were subsequently analysed with the computer program Data Explorer^{TM}, following the recommendation from the supplier, with help from Dr. Andrew Cronshaw.

ii) ICP-MS

Purified protein samples (1.5-2 ml), which were all later normalised to 60 µg/ml were analysed using an Agilent 7500ce (with octopole reaction system) instrument, employing a radio frequency (RF) forward power of 1540 W and reflected power of 1 W, with argon gas flows of 0.82 L/min and 0.2 L/min for carrier and makeup flows, respectively. Sample solutions were taken up into the Micro mist using a peristaltic pump at a rate of approximately 1.2 ml/min. Skimmer and sample cones were made of nickel. The instrument was operated in spectrum multi-tune acquisition

mode and five replicate runs per sample were employed. Each mass was analysed in fully quantitiative mode (three points per unit mass). The following masses were selected for analysis: ^{26}Mg, ^{55}Mn, ^{56}Fe, ^{57}Fe, ^{59}Co, ^{60}Ni, ^{63}Cu, ^{65}Cu, and ^{66}Zn. Two tune steps were used to analyse the metals: 'no gas' mode then 'helium mode' for ^{26}Mg, ^{56}Fe, ^{57}Fe, ^{60}Ni, ^{63}Cu, ^{65}Cu, and ^{66}Zn to remove any polyatomic interferences.

Calibration standards with a range of metal concentrations (0, 1, 10 and 100 µg/L) were prepared by sequential dilution from single element 1000 mg/L stock solutions (Fisher Scientific UK LTD Bishop Meadow Road, Loughborough, Leicestershire LE11 5RG) in 20 mM TEA-HCl buffer pH 7.6 and 50 mM NaCl, as well as in water to assess matrix effects. Samples were analysed using the standards prepared in 20 mM TEA-HCl buffer pH 7.6. Ten blank samples of the 20 mM TEA-HCl buffer pH 7.6 were analysed to calculate the LOD for each metal. Where values fell below the limit of detection they were quoted as a < value (Fuad et al. 2011). This analysis was performed by Dr. Lorna Eades.

2.3 Experimental for Chapter 5

2.3.1 Preparation of *T. brucei* cytosolic fractions

T. brucei cytosolic fractions were prepared by differential centrifugation as described previously, in isotonic buffer (250 mM sucrose (Fisher Scientific), 25 mM Tris-HCl (Sigma Life Sciences) pH7.4, 1 mM EDTA) (Misset and Opperdoes 1984). EDTA was included primarily as a proteolytic inhibitor, but is also a useful form of metal sequestration. It was not removed from the fractions, and would thus still be present with its bound metals during the ICP-OES measurements. Soluble proteins in the cytosolic fractions were prepared by Melisa-Gualdron Lopez and Muriel Mazet, de Duve Institute, Brussels, Belgium.

The concentrations of proteins in cytosolic samples were measured by Bradford assay (Bradford 1976) with Coomassie Plus Protein Assay Reagent (Thermo Scientific), and BSA as a standard (see Appendix I (a)). Seven samples (3 ml each) were obtained that had been prepared on different occasions: three cytosolic samples from *in* vitro cultured procyclic insect-stage parasites (S1, S2, S3), and four

cytosolic samples from pathogenic bloodstream-form parasites (S4, S5, S6 and S7). Samples S5 and S6 gave metal analysis results with relatively high standard deviations, and have been omitted. iPGAM enzymatic assays of each cytosolic fraction were done using a Multimode Plate Reader-Molecular Devices M5 instrument, controlled by SoftMax Pro software version 5.4, by the addition of a 10 µl sample of the cytosolic fraction to give a final reaction mixture of 100 µl containing 100 mM TEA-HCl buffer pH 7.6, 1.5 mM 3PGA, 5 mM $MgCl_2$, 50 mM KCl, 0.8 mM NADH, 1 mM ADP, 2 units of ENO, 4 units of PYK and 6 units of LDH. The decreased absorbance of NADH at A_{340} nm was used to obtain the rate of reaction for specific activity measurements (one unit corresponds to the conversion of 1 µmol of substrate min^{-1} mg^{-1} protein under standard conditions) (Fuad et al. 2011).

2.3.2 ICP-MS and ICP-OES

The experimental details for ICP-MS are given in section 2.2.7 ii because the analysis is also relevant for chapter 4. It is relevant to note that these analyses measure total metal concentrations in the samples, and are regardless of the metals oxidation state.

ICP-OES was used to analyse 2.5-3 ml of cytosolic fractions with a Perkin Elmer Optima 5300 DV instrument, using an RF forward power of 1400 W, with argon gas flows of 15, 0.2 and 0.75 L/min for plasma, auxiliary, and nebuliser flows, respectively. Using a peristaltic pump, the samples were taken up into a Gem Tip cross-Flow nebuliser and Scotts spray chamber at a rate of 1.50 ml/min. The instrument was operated in axial mode, and the selected wavelengths for each element were analysed in fully quantitative mode (three points per unit wavelength). The different cytosolic fractions were all normalised to 1mg/ml protein, and three replicate runs per sample were carried out. Initially two wavelengths were selected for each element, but after the analysis was completed the following wavelengths were chosen for reporting results: Co 228.616 nm, Cu 327.393 nm, Fe 238.204 nm, Mg 285.213 nm, Mn 257.610 nm, Ni 231.604 nm and Zn 206.200 nm.

To assess the limit of detection (LOD) for each metal and the matrix effects of the buffers, a range of calibration standards was prepared both in deionised water (18Ω, Elga USF) and in cytosol buffer (250 mM sucrose, 25 mM Tris-HCl, pH 7.4). Standards of 0, 0.1, 10 and 100 mg/L were prepared by sequential dilution from single element 1000 mg/L stock solutions (Fisher Scientific UK LTD Bishop Meadow Road, Loughborough, Leicestershire LE11 5RG). With all of the calibration results the correlation coefficients for the linear regression were 0.9992 or better. Ten blank samples of each buffer were analysed after the instrument was calibrated, and the LOD is given by the formula $2.26*SD_b/Slope$ (where b is the standard deviation of the 10 blanks). The limits of detection are shown in chapter 5.

For analysis of the samples, a range of calibration standards from 0-500 mg/L was prepared in cytosol buffer. The majority of the samples fell in the range 200-500 mg/L, and results were calculated against a calibration from 10-500 mg/L. However where values were lower than 20 mg/L, the samples were reintegrated against a calibration only extending to 10 mg/L. Where values fell below the limit of detection they were quoted as a < value. It is relevant to mention that attempts were initially made to analyse the cytosolic fractions by ICP-MS, but were found to be unsuitable because of technical problems (Fuad et al. 2011). This analysis was performed by Dr. Lorna Eades.

2.3.3 MESPEUS

The MESPEUS database offers information concerning the geometry of metal sites in protein structures (Hsin et al. 2008), and was accessed through the CTCB website with the link: http://mespeus.bch.ed.ac.uk/MESPEUS_10/_1.jsp.

2.4 Experimental for Chapter 6

2.4.1 Discontinuous coupling enzyme assay

The discontinuous assay was employed by incubating a sample of iPGAM (0.162 μM) with a final concentration of 50 mM 3PGA for 15 min at 25°C, after an earlier incubation of the enzyme with a final concentration of 1 mM of each of these metal

salts: $CaCl_2$, $CoCl_2$, $CuCl_2$, $FeCl_2$, $MgCl_2$, $MnCl_2$, $NiCl_2$ and $ZnCl_2$ (Sigma Life Sciences) for 1 h at the same temperature, in the presence of 200 mM TEA-HCl buffer pH 7.6 and 50 mM NaCl which was prepared for a maximum buffering capacity. The stock solutions of the enzyme components and the metal salts were prepared using 200 mM TEA-HCl pH 7.6. The reaction was quenched by incubating the mixture at 100°C for 2 min, before adding 5 µl samples from the first tubes into the 96 well-plates containing 95 µl of the reaction mixture (100 mM TEA-HCl buffer pH 7.6, 5 mM $MgCl_2$, 50 mM KCl, 1 mM NADH, 1 mM ADP, 1 unit of ENO, 2 units of PYK and 3 units of LDH. Controls were carried out with the addition of 2PGA into the reaction mixture, in the absence of *Lm*iPGAM and 3PGA. The oxidation of NADH to NAD^+ was observed as the end absorbance values, which was compared to the calibration curves in order to obtain the concentrations of 2PGA in the wells. The NADH conversion was monitored using the Multimode Plate Reader-Molecular Devices M5 instrument.

Prior to the analysis, a series of calibration curves was generated. These standard curves were obtained by experimentally adding known concentrations of 2PGA (0.63 mM, 0.31 mM, 0.16 mM, 0.08 mM and 0 mM) into the assay mixture containing the same components as stated above. Calibration standards were also generated in the presence of different metals, explicitly at three different metal concentrations: 0.05 mM, 0.1 mM and 0.5 mM.

2.5 Experimental for Chapter 7

2.5.1 Virtual screens programs: UFSRAT, AutoDock, and AutoDock Vina

The program UFSRAT was used to screen 3PGA/2PGA analogues from a multiconformer library comprising 4,853,000 molecules from various suppliers: Chembridge, Asinex, Maybridge, Enamine, LifeChemicals (InfLab) Specs, InterBioScreen and ChemDiv. The SDF files which represent the atom distributions of the query molecules were used as the input. This program can be accessed through the CTCB website with the link: http://opus.bch.ed.ac.uk/ufsrat/. The results obtained were listed based on the similarity scores and were saved as SDF files. Subsequently,

AutoDock and AutoDock Vina screens were performed by using the closed-form structure of *Lm*iPGAM and 2PGA as the input structure and original ligand, respectively. The screen was conducted against the SDF files comprising the previous UFSRAT results, and were listed based on the AutoDock and AutoDock Vina scores. The COmbining Docking and Similarity Search (CODASS) analysis was performed by Dr. Douglas Houston.

2.5.2 The quantitative high-throughput screening (qHTS)

The qHTS was conducted in the NIH, Bethesda, USA. The screening was performed using the Kinase-Glo® Luminescent Kinase assay method which quantifies the amount of ATP remaining in the sample following a kinase reaction (Promega). The qHTS screen for *Lm*iPGAM was employed initially with the Library of Pharmacologically Active Compounds (LOPAC®) (Sigma-Aldrich).

2.5.3 Inhibition analysis with selected compounds

The inhibition analysis was performed using both the continuous and discontinuous coupling enzyme assay approaches, following the same conditions as stated in sections 2.2.2 and 2.4.1, respectively, except for the incorporation of an additional step: the incubation with the selected compounds. Both assays were performed by incubating a final concentration of 0.01 mg/ml *Lm*iPGAM from the ion-exchange column with a final concentration of 1 mM Co^{2+} at 25°C for 1 h, before adding a final concentration of 1 mM compounds individually with further incubation for another hour. The remaining steps were as described in the previous 2.2.2 and 2.4.1 sections. It is noteworthy that the initial stock solutions of the compounds were prepared by dissolving them with 100% DMSO except for compound C5 and C10 (50% DMSO), before being further diluted to prepare a 10 mM stock solution with 200 mM TEA-HCl pH 7.6. The oxidation of NADH to NAD^+ was monitored using the Multimode Plate Reader-Molecular Devices M5 instrument.

Chapter 3:

Development of Protein Expression Systems and New Purification Approaches for Trypanosomatid iPGAMs

3.1 Aims

Trypanosomatid iPGAMs, as members of the alkaline phosphatase superfamily, require metal ions to support their integrity and catalytic function (Chevalier et al. 2000; Guerra et al. 2004; Poonperm 2005; Nowicki et al. 2009). Co^{2+} was shown to stabilise *Lm*iPGAM (Poonperm 2005), and has often been included in the whole protein characterisation process: from enzyme purification to the crystallisation stage. The closed-form crystal structure of *Lm*iPGAM (Poonperm 2005; Nowicki et al. 2009) as well as the open-form *Tb*iPGAM (Mercaldi et al. 2012) suggest that Co^{2+} is in the active site, indicating that Co^{2+} has high affinity to the enzyme. In fact, *Lm*iPGAM has always been purified in the presence of Co^{2+}, and the metal is required for activity measurements and crystallisation conditions too. However, these are *in vitro* conditions, and similar circumstances are unlikely to occur in *vivo*, where the Co^{2+} concentration may not be present at a sufficient concentration to support iPGAM activity.

An investigation into the roles played by Co^{2+} in the structure and function of *Lm*iPGAM requires that the enzyme be purified in the absence of added Co^{2+}. Thus, for the purpose of eliminating the presence of Co^{2+} in the whole process *in vitro*, assuming that there is no residual metal obtained from the bacterial cells during expression, the purification steps have to be employed in its absence. This would necessitate either: i) avoiding the use of a Talon metal-affinity column, or ii) replacing Co^{2+} with other metals, for example Ni^{2+} during the metal-affinity chromatography step. Apart from preventing Co^{2+} contamination, another important aspect in *Lm*iPGAM analysis is to direct the enzyme conformation into the open-form state which is important for the purpose of drug design, as well as a completion to the previous closed-form structure. In this case, additional chromatography steps would be required, such as ion-exchange and gel filtration steps, which have the ability to separate proteins based on the charge distribution on the surface of the

folded proteins, as well as by shape and size, respectively. Hence, in this chapter, a new purification system for iPGAM will be discussed in detail.

It is noteworthy that in this study, only the His-tagged versions of *Lm*iPGAM and *Tb*iPGAM were expressed and purified. While the un-tagged version of *Tb*iPGAM was successfully overexpressed, neither the un-tagged versions of *Lm*iPGAM nor *Tb*iPGAM were purified. This is due to the efficiency of the His-tagged purification protocol in yielding high purity protein samples.

3.2 Introduction

Protein purification involves a number of important stages which begin with a disruption of the cells that contain the specific protein of interest, usually bacteria, insect, yeast or mammalian cells (Marino 1989; Makrides 1996). Overexpression of soluble protein is normally required to ensure maximal protein production can be achieved through purification. The starting material however, does not only contain the intended protein, but other macromolecules such as DNA and RNA, as well as a mixture of polysaccharides, lipids, smaller molecules as well as other proteins (Burgess 2008). It is vital thus, to disrupt the particular cells to release the protein, and therefore enable successful separation of the protein of interest from the other macromolecules, including other proteins. Often the explicit characteristics of a specific protein are taken into account before any purification procedure is chosen, for example, whether the intended protein possesses distinct structural properties that are distinguishable from the others. Chromatography is one of the techniques utilised for this purpose, in which proteins in the mobile phase may be separated according to their ability to adsorb to a matrix (stationary phase) (Price and Nairn 2009). The interaction between the protein and the matrix varies, and the matrix is specifically chosen based on the protein's properties. For example, ion-exchange and hydrophobic interaction chromatography are based on the surface charge of the protein, gel filtration depends on size, and affinity relies on specific binding sites in the protein structure (Price and Nairn 2009).

While chromatography is an excellent technique for protein purification, it is important to note that the incorporation of "non-native" metals, especially in metalloproteins, may occur significantly. This has been shown in some proteins, for example the expression of *Pseudomonas aeruginosa* azurin in *E. coli* has misleadingly introduced Zn^{2+}, instead of Cu^{2+}, the native metal for the protein (Nar et al. 1992). It is important thus, to carefully select the expression and purification system so as to minimise the introduction of non-biologically relevant metals in the protein structure. Otherwise, the interpretation from the enzymatic activity or crystal structure may provide a misleading understanding.

Over the years, purification of *Lm*iPGAM and *Tb*iPGAM has frequently been conducted by a TALON metal-affinity resin (Chevalier et al. 2000; Guerra et al. 2004; Poonperm 2005), except for a Ni^{2+}-NTA agarose column in a study by Collet et al. 2001. The separation is based on the interaction between a His-tag that is attached to the specific protein, with the affinity column containing the bound metal. For *Lm*iPGAM in particular, Co^{2+} was chosen as the bound metal in the resin, which was reported to be the only metal that could maximally reactivate iPGAM activity after metal-chelating experiments (Guerra et al. 2004). It is known that Co^{2+} has higher selectivity for His-tags compared to other metals (http://www.gbiosciences.com/ResearchProducts/Cobalt-Chelating-Resin.aspx), thus providing a reason for choosing Co^{2+} over other metals, such as Ni^{2+}. The issue of which metal ions support iPGAM activity in the parasite cell, based on the knowledge that Co^{2+} concentrations in eukaryotic cells are very low (Finney and O'Halloran 2003; Maret 2010), raised the concern that Co^{2+} is not the biologically relevant metal for the trypanosomatid iPGAMs. Hence, for the purpose of preventing Co^{2+} from being trapped in the active site of iPGAM during the whole purification process, Ni^{2+} would be utilised in the metal-affinity chromatography step, and any initial enzymatic activity would be measured in the absence of added Co^{2+}. A high sequence similarity between *Lm*iPGAM and *Tb*iPGAM (Figure 3.1) suggested that a similar procedure could be performed in purifying both enzymes. By combining this procedure with metal analysis studies which will be discussed in detail in chapter 5 and 6, the biologically relevant metal for trypanosomatid iPGAMs may be

unravelled. The amino acid sequences of *Tb*iPGAM and *Lm*iPGAM which have been used as constructs in the experiments are shown in Figure 3.2.

```
Tb  -MALTLAAHKTLPRRKLVLVVLDGVGIGPRDEYDAVHVAKTPLMDALFND 49
Lm  MSALLLKPHKDLPRRTVLIVVMDGLGIGPEDDYDAVHMASTPFMDAHRRD 50
      ** * .** ****.:::**:**:****.*:*****:*.**:***  .*

Tb  PKHFRSICAHGTAVGLPTDADMGNSEVGHNALGAGRVVLQGASLVDDALE 99
Lm  NRHFRCVRAHGTAVGLPTDADMGNSEVGHNALGAGRVALQGASLVDDAIK 100
     :***.: ****************************.**********::

Tb  SGEIFTSEGYRYLHGAFSQPGRTLHLIGLLSDGGVHSRDNQVYQILKHAG 149
Lm  SGEIYTGEGYRYLHGAFSKEGSTLHLIGLLSDGGVHSRDNQIYSIIEHAV 150
    ****:*.***********: * ******************:*.*::**

Tb  ANGAKRIRVHALYDGRDVPDKTSFKFTDELEEVLAKLREGGCDARIASGG 199
Lm  KDGAKRIRVHALYDGRDVPDGSSFRFTDELEAVLAKVRQNGCDAAIASGG 200
     :****************** :**:****** ****:*:.**** *****

Tb  GRMFVTMDRYEADWSIVERGWRAQVLGEGRAFKSAREALTKFREEDANIS 249
Lm  GRMFVTMDRYDADWSIVERGWRAQVLGDARHFHSAKEAITTFREEDPKVT 250
    **********:****************:.* *:**:**:*.*****.:::

Tb  DQYYPPFVIAGDDGRPIGTIEDGDAVLCFNFRGDRVIEMSRAFEEEEFDK 299
Lm  DQYYPPFIVVDEQDKPLGTIEDGDAVLCVNFRGDRVIEMTRAFEDEDFNK 300
    *******::..::.:*:***********.**********:****:*:*:*

Tb  FNRVRLPKVRYAGMMRYDGDLGIPNNFLVPPPKLTRTSEEYLIGSGCNIF 349
Lm  FDRVRVPKVRYAGMMRYDGDLGIPNNFLVPPPKLTRVSEEYLCGSGLNIF 350
    *:***:*******************************.***** *** ***

Tb  ALSETQKFGHVTYFWNGNRSGKLSEERETFCEIPSDRVQFNQKPLMKSKE 399
Lm  ACSETQKFGHVTYFWNGNRSGKIDEKHETFKEVPSDRVQFNEKPRMQSAA 400
    * ********************:.*::*** *:********:** *:*

Tb  ITDAAVDAIKSGKYDMIRINYPNGDMVGHTGDLKATITSLEAVDQSLQRL 449
Lm  ITEAAIEALKSGMYNVVRINFPNGDMVGHTGDLKATITGVEAVDESLAKL 450
    **:**::*:*** *:::***:*****************.:****:** :*

Tb  KEAVDSVNGVFLITADHGNSDDMVQRDKKGKPVRDAEGNLMPLTSHTLAP 499
Lm  KDAVDSVNGVYIVTADHGNSDDMAQRDKKGKPMKDGNGNVLPLTSHTLSP 500
    *:********:::**********.********::*.:**::*******:*

Tb  VPVFIGGAGLDPRVQMRTDLPRAGLANVTATFINLMGFEAPSDYEPSLIE 549
Lm  VPVFIGGAGLDPRVAMRTDLPAAGLANVTATFINLLGFEAPEDYEPSLIY 550
    ************** ****** *************:*****.*******

Tb  VA- 551
Lm  VEK 553
    *
```

Figure 3.1 Sequence alignment of *Lm*iPGAM and *Tb*iPGAM resulted in 78% sequence identity. Amino acids with similar properties are shown with specific colours. The red colour corresponds to small and hydrophobic amino acids, the pink corresponds to basic amino acids, the blue corresponds to acidic amino acids and the green corresponds to hydroxyl and amine amino acids. The symbols under the residues indicate "*" as identical residues; ":" as semi-conserved residues; and "." as low conservation. The analysis was done using the sequence alignment program CLUSTALW2 (http://www. ebi.ac.uk/ Tools/msa/clustalw2/), with the parameters set at default values.

a

His-tagged *Tb*iPGAM

MGSSHHHHHH SSGLVPRGSH MALTLAAHKT LPRRKLVLVV LDGVGIGPRD EYDAVHVAKT

PLMDALFNDP KHFRSICAHG TAVGLPTDAD MGNSEVGHNA LGAGRVVLQG ASLVDDALES

GEIFTSEGYR YLHGAFSQPG RTLHLIGLLS DGGVHSRDNQ VYQILKHAGA NGAKRIRVHA

LYDGRDVPDK TSFKFTDELE EVLAKLREGG CDARIASGGG RMFVTMDRYE ADWSIVERGW

RAQVLGEGRA FKSAREALTK FREEDANISD QYYPPFVIAG DDGRPIGTIE DGDAVLCFNF

RGDRVIEMSR AFEEEEFDKF NRVRLPKVRY AGMMRYDGDL GIPNNFLVPP PKLTRTSEEY

LIGSGCNIFA LSETQKFGHV TYFWNGNRSG KLSEERETFC EIPSDRVQFN QKPLMKSKEI

TDAAVDAIKS GKYDMIRINY PNGDMVGHTG DLKATITSLE AVDQSLQRLK EAVDSVNGVF

LITADHGNSD DMVQRDKKGK PVRDAEGNLM PLTSHTLAPV LFLSEALVLI PVCKCGQTFR

VRPCNVTATF INLMGFEAPS DYEPSLIEVA

b

His-tagged *Lm*iPGAM

MSALLLKPHK DLPRRTVLIV VMDGLGIGPE DDYDAVHMAS TPFMDAHRRD NRHFRCVRAH

GTAVGLPTDA DMGNSEVGHN ALGAGRVALQ GASLVDDAIK SGEIYTGEGY RYLHGAFSKE

GSTLHLIGLL SDGGVHSRDN QIYSIIEHAV KDGAKRIRVH ALYDGRDVPD GSSFRFTDEL

EAVLAKVRQN GCDAAIASGG GRMFVTMDRY DADWSIVERG WRAQVLGDAR HFHSAKEAIT

TFREEDPKVT DQYYPPFIVV DEQDKPLGTI EDGDAVLCVN FRGDRVIEMT RAFEDEDFNK

FDRVRVPKVR YAGMMRYDGD LGIPNNFLVP PPKLTRVSEE YLCGSGLNIF ACSETQKFGH

VTYFWNGNRS GKIDEKHETF KEVPSDRVQF NEKPRMQSAA ITEAAIEALK SGMYNVVRIN

FPNGDMVGHT GDLKATITGV EAVDESLAKL KDAVDSVNGV YIVTADHGNS DDMAQRDKKG

KPMKDGNGNV LPLTSHTLSP VPVFIGGAGL DPRVAMRTDL PAAGLANVTA TFINLLGFEA

PEDYEPSLIY VEKLEHHHHH H

Figure 3.2 a) The amino acid sequence of *Tb*iPGAM including the His-tag. b) The amino acid sequence of *Lm*iPGAM including the His-tag. The first and the last residues of the native sequences are highlighted in yellow.

3.3 The expression systems for *Lm*iPGAM

The bacterially expressed trypanosomatid iPGAMs were initially proposed to be purified in the presence and absence of His-tags. Both the His-tagged *Lm*iPGAM and *Tb*iPGAM plasmids were available in the laboratory, and had been obtained from Prof. Paul Michels, De Duve Institute, Brussels. The genes were present in pET28a plasmids, and were carried by *E. coli* strain DH5α. The construction of both the untagged *Lm*iPGAM and *Tb*iPGAM genes was employed to produce plasmids suitable for purification of the enzymes in the absence of metal-affinity chromatography. The general molecular cloning techniques were described in chapter 2, and the steps involved in the development of the expression systems will be discussed in detail in the next sections. However, for the purpose of purification, only the His-tagged versions of the protein were purified. This is due to the efficiency of the protocol to yield high purity protein samples.

i) Histine-tagged *Lm*iPGAM expression system

The His-tagged *Lm*iPGAM expression system was a gift from Prof. Paul Michels and has been developed in the Brussels research group. The 6His-tag was present at the C-terminus of the protein (LEHHHHHH), and the molecular mass was calculated to be 60723.3 Da (excluding the His-tag and including the initiator methionine), with a theoretical pI of 5.43. Since the protein was expressed in *E. coli*, the initiator methionine was translated and included in the mass calculation, but the formyl group was normally removed after protein translation (Sherman et al. 1985; Spector et al. 2003). Meanwhile, when the His-tag was included in the calculation, the molecular mass increased to 61788.5 Da, with a theoretical pI of 5.62. The protein parameters, which were calculated using the Expasy ProtParam (http://web.expasy.org/cgi-bin/protparam/protparam) are tabulated in Table 3.1. In the subsequent purification experiment (see section 3.4i), the plasmid encoding the His-tag protein was transformed into *E. coli* strain BL21 (DE3) before protein expression following standard conditions as described by Poonperm 2005. The purification steps will be discussed in detail in section 3.4 i.

Table 3.1 *Lm*iPGAM parameters obtained from the Expasy database.

Protein	Molecular mass (Da)	Theoretical pI	Number of amino acids	Extinction coefficient (M^{-1} cm^{-1}) *
His-tagged *Lm*iPGAM	61788.5	5.62	561	42080
Un-tagged *Lm*iPGAM	60723.3	5.43	553	42080

* As measured in water at A_{280} nm

ii) Un-tagged *Lm*iPGAM expression system

The un-tagged *Lm*-iPGAM expression plasmid was initially developed to eliminate the presence of the His-tag in the construct. As mentioned earlier, the *Lm*iPGAM gene had been originally inserted into the pET28a vector, which carries an N-terminal His-tag, thrombin cleavage site and a T7 promoter, with an optional C-terminal His-tag sequence (Novagen®). Since the *Lm*iPGAM gene was fused with the C-terminal His-tag, two restriction sites in the multiple cloning site (MCS), *Nco*I and *Xho*I (Figure 3.3) were used to excise the iPGAM gene, hence leaving the C-terminal His-tag in the original pET28a vector. The next step involved PCR amplification of the target gene (*Lm*iPGAM) before being inserted into an intermediate vector (pJET1.2/blunt). In this case, a set of primers has been designed (see chapter 2), where the same restriction sites were added directly to the 5' and 3' ends of the primers. The resulting amplified PCR product with approximately 1.6 kb in size (Figure 3.4 a) was subsequently ligated into pJET1.2/blunt, which is a highly-efficient PCR-cloning vector, with the ability to insert either phosphorylated or non-phosphorylated DNA fragments (Fermentas). The success of the cloning procedure was evaluated by double digesting the clones with *Nco*I and *Xho*I, which clearly indicates the separation of the vector pJET1.2/blunt (2.95 kb) and *Lm*iPGAM insert (1.6 kb) (Figure 3.4 b). The final molecular cloning step was performed by insertion of the target gene from the intermediate vector into the final pET28a, with the utilisation of the same enzymes. The successful insertion of the *Lm*iPGAM gene was determined with a similar strategy as mentioned above, by double digesting with

both *Nco*I and *Xho*I, as well as a single digestion with the enzyme *Nco*I. The resulting bands (Figure 3.4 c) confirmed the attainment of the *Lm*iPGAM gene integration into pET28a. The constructed plasmid map is shown in Figure 3.5, and the correct sequence was later confirmed by Dundee Sequencing Service, the University of Dundee.

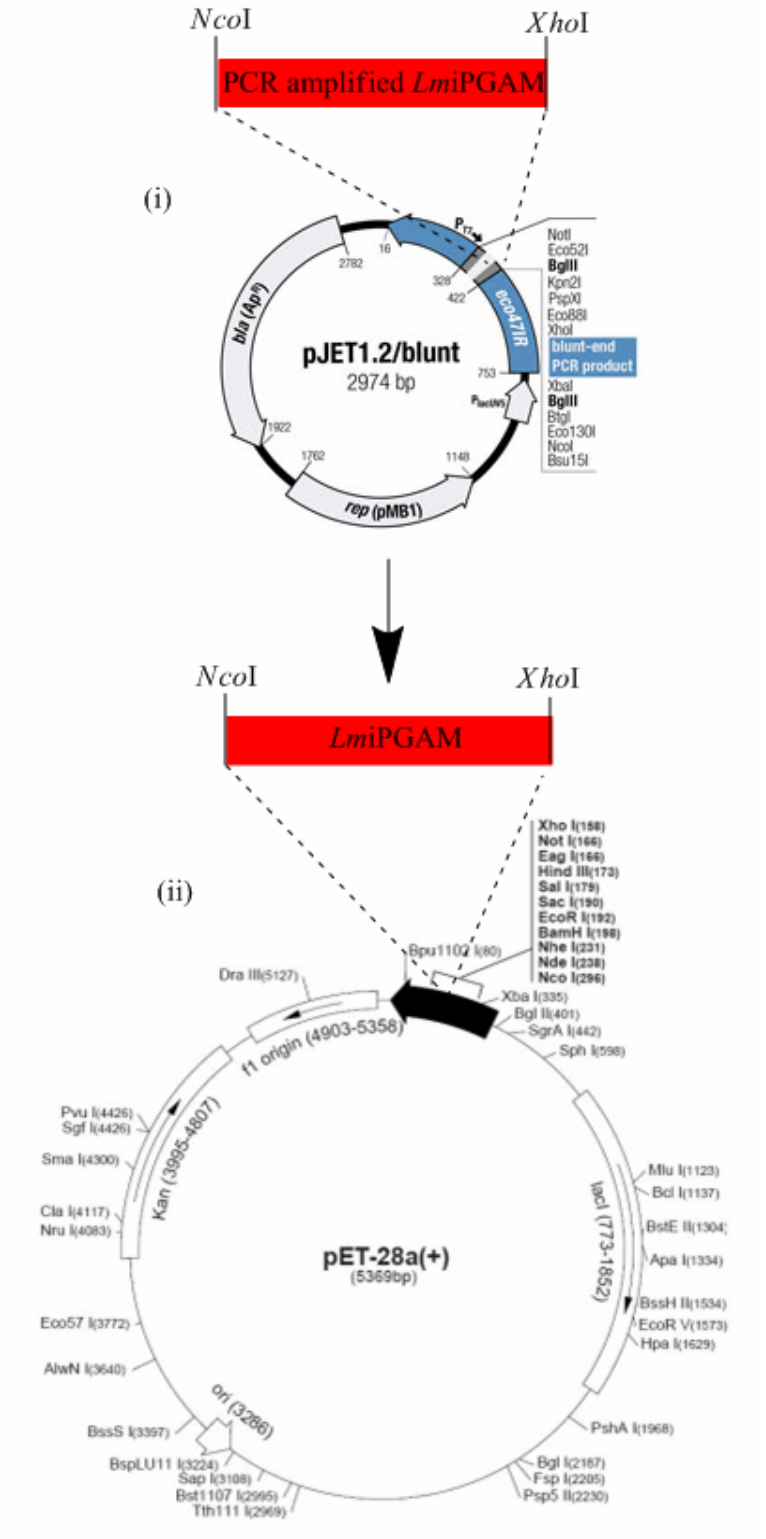


Figure 3.3 The cloning strategy for *Lmi*PGAM which involved two separate steps: (i) the insertion of the *Lmi*PGAM PCR product into an intermediate vector (pJET1.2/blunt), and (ii) the insertion of the *Lmi*PGAM gene from the intermediate vector into the final pET28a vector.

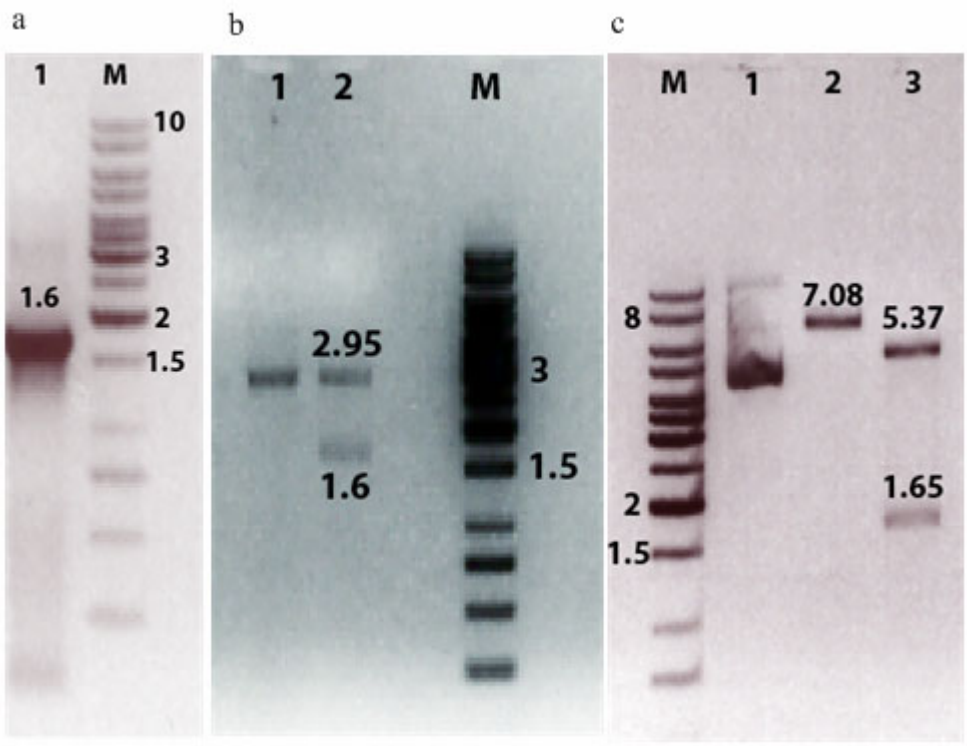


Figure 3.4 Agarose electrophoresis (0.8% (w/v)) of the PCR amplified *Lm*iPGAM and the digestion results. a) PCR amplified *Lm*iPGAM (well 1). b) The pJET1.2/blunt containing the undigested *Lm*iPGAM (well 1), and the pJET1.2/blunt containing digested *Lm*iPGAM with *Nco*I and *Xho*I (well 2). c) The undigested pET28a containing the untagged *Lm*iPGAM (well 1), recombinant clone single-digested with *Nco*I (well 2), and recombinant clone double-digested with *Nco*I and *Xho*I (well 3). GeneRuler^{TM} 1kb DNA ladder are shown as M.

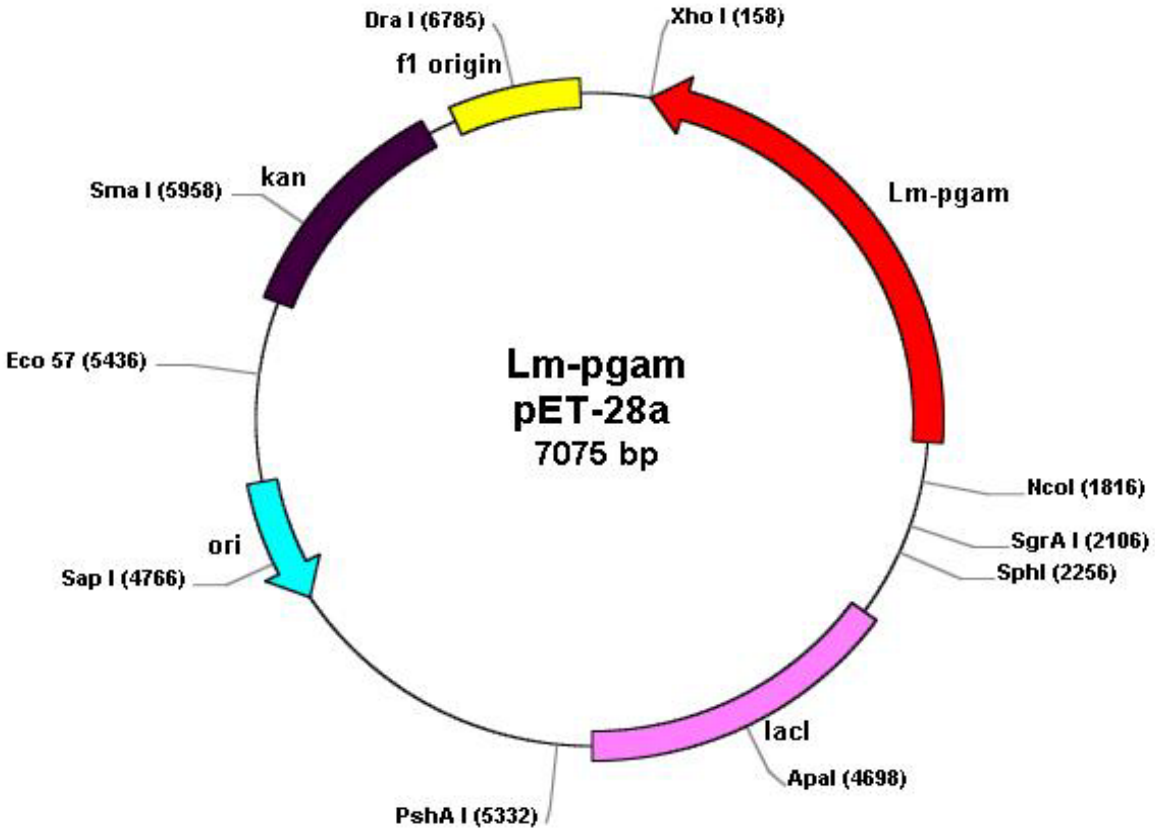


Figure 3.5 The plasmid map of the constructed un-tagged *Lm*iPGAM.

3.4 Purification steps for His-tagged *Lm*iPGAM

In the previous literature (Guerra et al. 2004; Poonperm 2005), *Lm*iPGAM was purified with a two-step procedure, involving a metal-affinity chromatography step using a Co^{2+}-containing TALON column, followed by a desalting/buffer-exchange step. The present study in contrast, involved three main purification steps for *Lm*iPGAM: immobilised metal-affinity chromatography, desalting and ion-exchange. It is noteworthy that the expression procedure was performed following the method described by Poonperm 2005, which produced a substantial amount of over-expressed protein (Figure 3.6 b). It is also important to note that Co^{2+} was not added in any purification steps or buffer preparation (see chapter 2). The un-tagged *Lm*iPGAM expression trials (see Appendix I (b)) were unsuccessful in obtaining

soluble *Lm*iPGAM for protein purification, thus it was decided that the purification procedure would only focus on the His-tagged *Lm*iPGAM.

i) Step one: immobillised metal-affinity chromatography

The utilisation of polyhistidine-tags in proteins for chromatographic purposes has been widely reported (Hemdan 1989; Porath 1992; Gaberc-Porekar and Menart 2001). The principle of the technique lies on the interaction between metal ions that are entrapped in chromatographic supports containing covalently bound chelating compounds (Gaberc-Porekar and Menart 2001) with specific amino acid residues which are exposed on the surface of the protein molecule. One of the most attractive features for this chromatographic step is the ability to purify a large amount of protein during a rapid purification process. The specificity of the technique to purify explicit proteins however, is considered to be moderate (Gaberc-Porekar and Menart 2001), hence frequently requiring further additional purification procedures. Nevertheless, it has evolved as one of the most efficient steps in a chromatographic procedure for larger scale-ups.

In this study, the chromatographic step was employed through immobilisation of Ni^{2+} ions in a column containing a Sepharose 6 matrix, which is packed with highly cross-linked agarose beads with a covalently immobilised chelating group (GE Healthcare). The cell lysate containing *Lm*iPGAM flowed over the 5 ml Fast Flow IMAC column (GE Healthcare), where the specifically engineered His-tagged protein bound to the stationary phase through the interactions between imidazole side chains of histidines and the Ni^{2+} ions, while any unbound molecules were washed away as the flow-through. The elution was achieved in the presence of increasing imidazole concentration, which involved two separate stages, known as the washing and elution steps. In the initial washing step, 20 mM imidazole was included in the buffer (see chapter 2) with His-tagged protein bound to the solid matrix, while any other weakly-bound proteins, which are contaminants from the bacterial culture pass through the column (Figure 3.6 a). Surprisingly however, SDS-PAGE analysis (Figure 3.6 b) showed that the peak corresponding to the washing contaminants also contains a significant presence of iPGAM, which suggests that the washing could be

done at a lower imidazole concentration. After an additional 20 CV wash at 5% buffer B, the elution step was performed in a step-wise fashion, from 25 mM to 500 mM imidazole. A peak was detected by UV sensor at A_{280} nm (Figure 3.6 a) which corresponds to *Lm*iPGAM (Figure 3.6 b). The total yield of the protein from 1L of culture was found to be 20 mg.

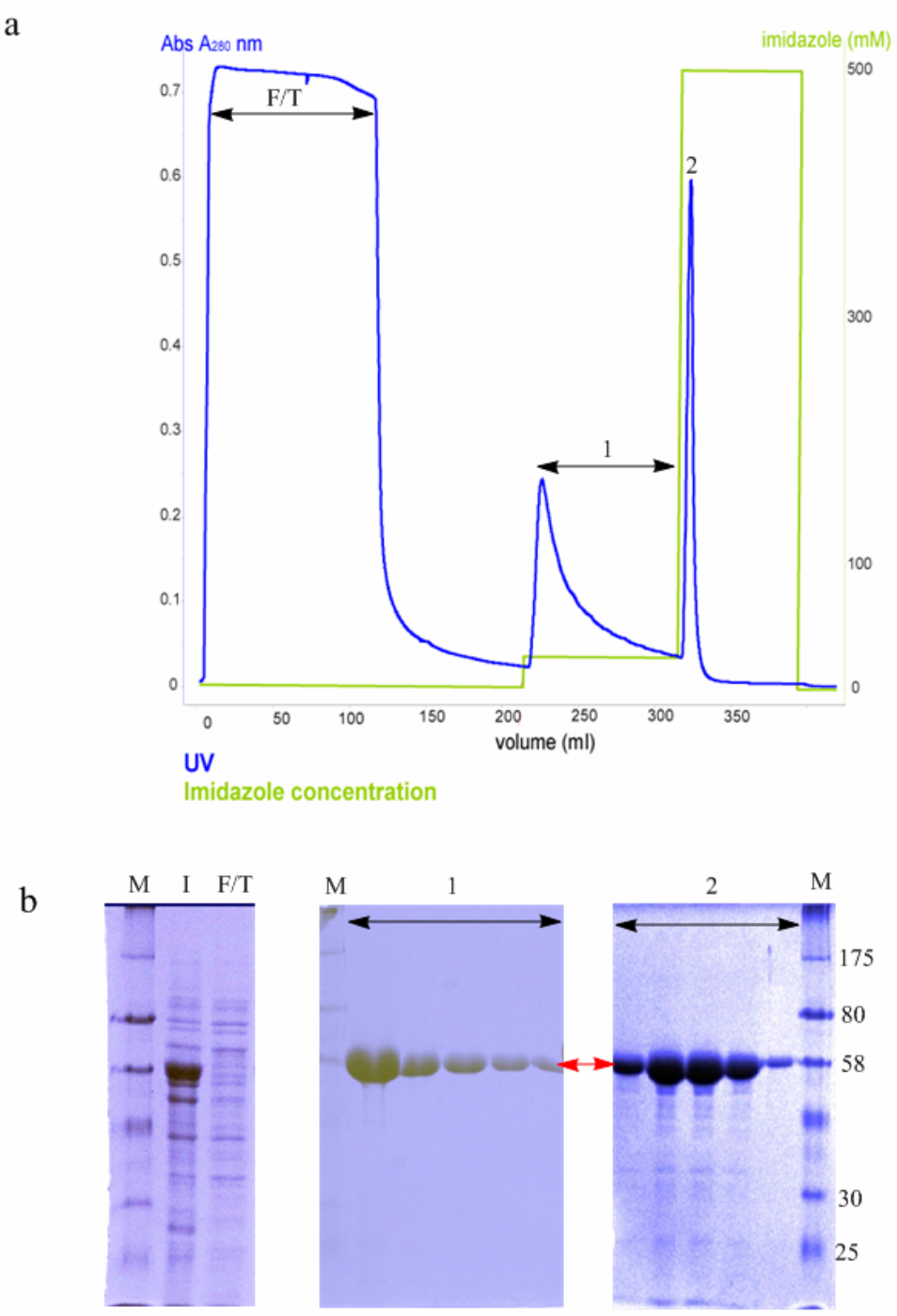


Figure 3.6 a) The elution profile of *Lmi*PGAM from 5 ml Fast Flow IMAC column resulting in two distinct elution sections (1 and 2) as well as an earlier flow-through (F/T) section (blue trace) b) A 12% SDS-PAGE is shown with 2-4 µg samples of each protein fraction with the initial extract (I), flow-through (F/T) and fractions from sections 1 and 2. A red arrow shows that the pure protein corresponds to the size of His-tagged *Lmi*PGAM (61.8 kDa). The protein markers are shown as M and labelled as kDa.

ii) Step two: desalting

The IMAC step is advantageous for rapid and large-scale protein purification, but nevertheless, also contributes to significant concentrations of buffer salts and imidazole introduced into the protein samples. This is a direct consequence of the high ionic strength required during the adsorption and elution steps in the chromatography procedure (Price and Nairn 2009). Often biophysical techniques are not compatible with high salt concentrations, thus requiring the samples to be exchanged into low salt environment. In the case of *Lm*iPGAM, salt removal was vital to provide a low ionic strength environment as a preparation for the subsequent chromatographic procedure, ion-exchange chromatography. It is important to note however, that Na^+ is always included through the whole purification process (50mM -300 mM at different steps of purification, see chapter 2) as it was found to stabilise *Lm*iPGAM (Poonperm 2005). The significant difference between the size of proteins and salts or smaller molecules made it plausible for the separation to be employed successfully. The principle of the technique is the same as size-exclusion chromatography (see chapter 4): higher molecular weight substances are unable to enter the gel pores and elute earlier, while lower molecular weight substances have the ability to enter the gel pores, resulting in a delayed elution (GE Healthcare).

Pure *Lm*iPGAM was passed through a HiPrep 26/10 desalting column, which is a pre-packed Sephadex G-25 Fine column. 20 mM HEPES buffer was used during column equilibration prior to sample application, as explained in detail in chapter 2. As depicted in Figure 3.7 a, a sharp UV peak at A_{280} nm was observed, indicating the successful elution of the protein of interest (*Lm*iPGAM). Subsequently, the conductivity signal rose as an indication of salts or any other smaller molecules eluted from the column. The fractions corresponding to the UV peak were analysed by SDS-PAGE analysis (Figure 3.7 b), where a single prominent band was exhibited and corresponds to the size of *Lm*iPGAM.

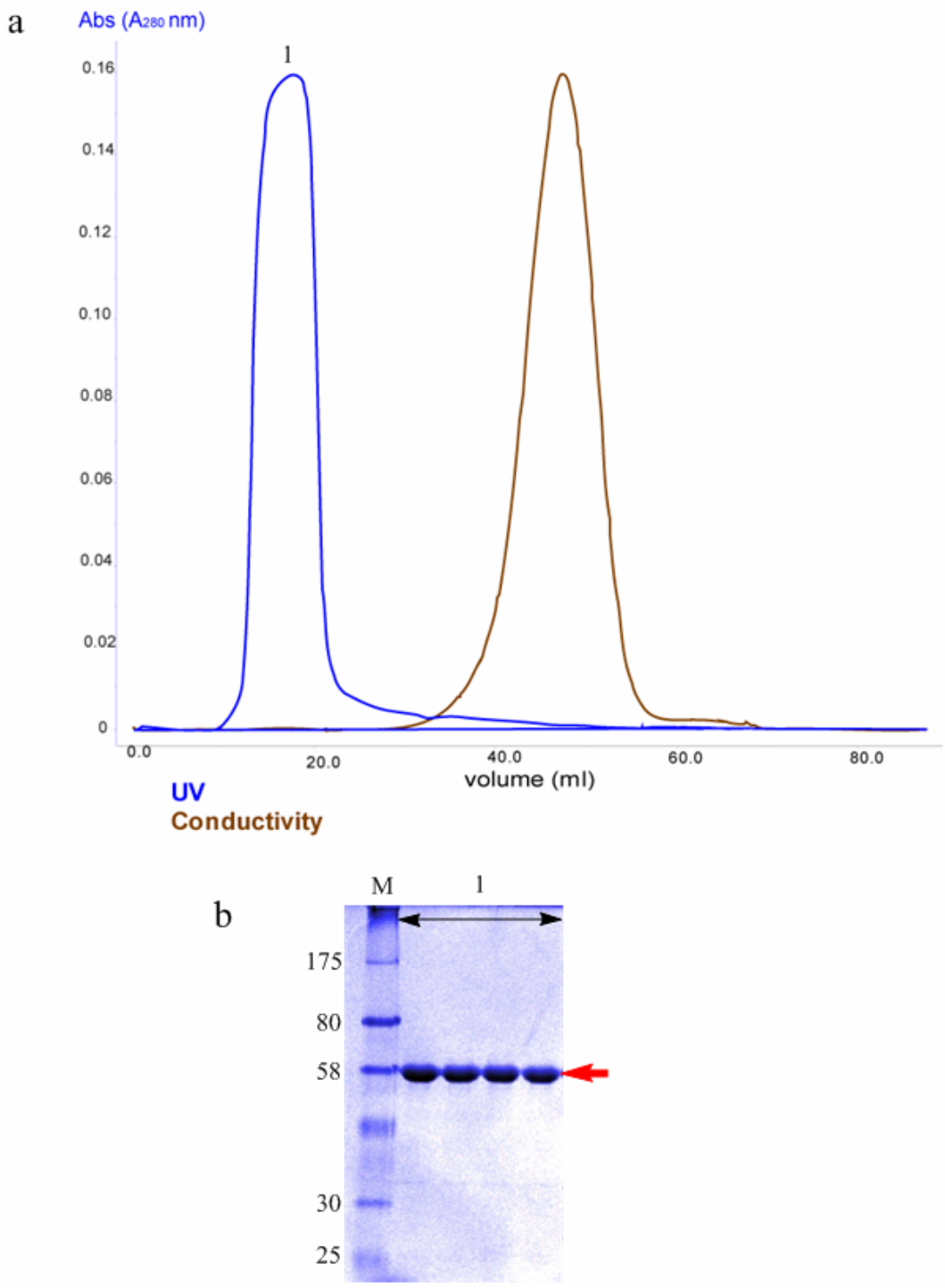


Figure 3.7 a) The elution profile of *Lm*iPGAM from the HiPrep 26/10 desalting column gave a single protein elution (blue trace), followed by salt and smaller molecules (brown trace) b) A 12% SDS-PAGE is shown with 2-4 µg samples of each protein fraction corresponding to the blue trace. The red arrow shows that the pure proteins correspond to the size of His-tagged *Lm*iPGAM (61.8 kDa). The protein marker is shown as M and labelled as kDa.

iii) Step three: ion-exchange chromatography

Different surface charges on the protein molecules contribute to the ability of proteins to be separated based on their surface charges through electrostatic interactions with a charged matrix (Price and Nairn 2009). In the analysis of *Lm*iPGAM, anion-exchange chromatography was employed to observe the interactions between the positively charged matrix and the protein molecules, which subsequently resulted in elution of proteins with theoretically distinct surface charges. The purification method can also be conducted as a polishing step to obtain higher purity protein samples.

MonoQ 5/50 GL is an anionic-exchange pre-packed glass column with monodispersed 10-µm porous beads giving high performance ion-exchange chromatography of proteins and other biomolecules (GE Healthcare). Protein samples in a low ionic strength environment, with a pH value that generates opposite net charges on the surface of the protein molecules from those of the column, result in binding between the protein and the matrix. By increasing the ionic strength of the elution buffer or changing the pH of the solution, the intended protein can be eluted. As MonoQ 5/50 GL is an anionic column, negatively-charged molecules would strongly bind to the matrix, whereas positively-charged molecules would not bind. In this experiment, a shallow salt gradient was obtained by increasing the concentration of NaCl to a maximum of 1 M. The resulting protein peaks which absorbed at A_{280} nm are illustrated in Figure 3.8 a.

Three distinct UV peaks were observed, exhibiting a reproducible elution pattern, as observed in different purification batches. SDS-PAGE analysis was employed to identify whether the fractions correspond to *Lm*iPGAM size (61.8 kDa), and it can be seen that all peak fractions had a similar molecular mass (Figure 3.8 b). These observations further augmented the possibility that *Lm*iPGAM, may well exist in different forms or conformations in solution, causing the charge distribution on the surface of the protein molecules to vary, which will be discussed further in chapter 4. Nevertheless, further analysis was required, to observe whether the differences were due to distinctive charge distributions, and not to incorrectly-folded protein or any

post-translational modification events that may have occurred. Further characterisation of the three peaks P1, P2 and P3 will be discussed in detail in the next chapter.

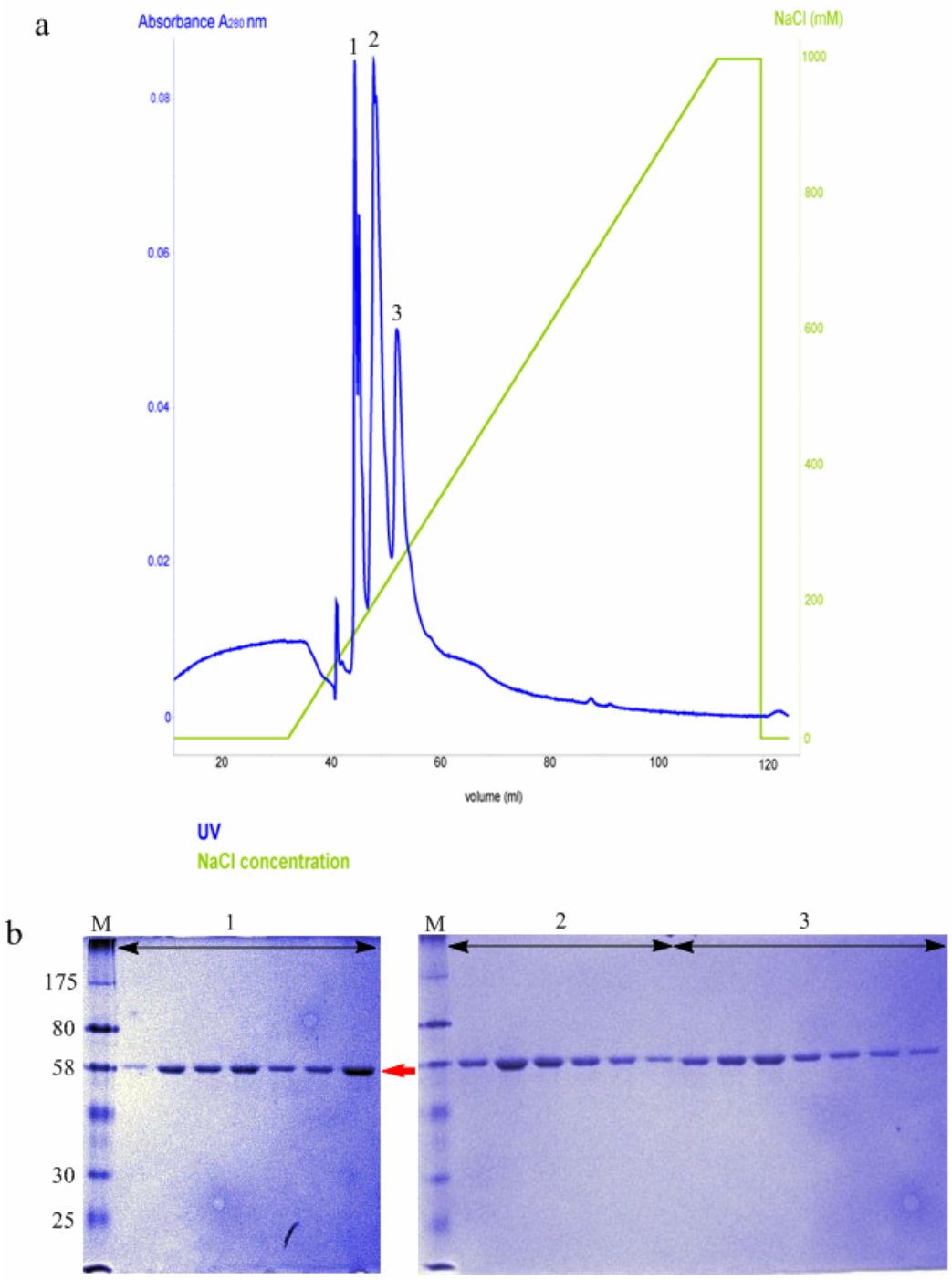


Figure 3.8 a) The elution profile of *Lmi*PGAM from MonoQ 5/50 GL column resulted in three separate peaks (blue trace) labelled 1, 2 and 3. b) A 12% SDS-PAGE is shown with 2-4 µg samples of each protein fraction corresponding to the three protein peaks (blue trace). The red arrow shows the pure proteins corresponding to the size of His-tagged *Lmi*PGAM (61.8 kDa). The protein markers are shown as M and labelled as kDa.

3.5 The expression systems for *Tb*iPGAM

Similar strategies were involved in the development of expression systems for *Tb*iPGAM, as explained for *Lm*iPGAM earlier in section 3.3. An N-terminal His-tagged *Tb*iPGAM expression plasmid was available in the laboratory, while the un-tagged *Tb*iPGAM necessitated the construction of a plasmid *via* molecular cloning. The plasmid pET28a containing the *Tb*iPGAM gene with an N-terminal His-tag was carried by *E. coli* strain DH5α. Subsequent molecular cloning techniques were employed for the purpose of the insertion of the *Tb*iPGAM gene with the His-tag eliminated.

i) His-tagged *Tb*iPGAM expression system

As mentioned in the earlier section (see section 3.3 i), the His-tagged *Tb*iPGAM expression system was a gift from Prof. Paul Michels and had been developed in Brussels research group. In contrast to *Lm*iPGAM, the 6His-tag was present at the N-terminus of the protein, and the His-tagged *Tb*iPGAM had a calculated molecular mass of 62852.1 Da with a theoretical pI of 5.98. However, when the His-tag was not included in the calculation, the molecular mass decreased to 60603.5 Da, with a theoretical pI of 5.56. The initiator methionine was always included in the calculation. The protein parameters, which were calculated using the Expasy ProtParam (http://web.expasy.org/cgi-bin/protparam/protparam) are tabulated in Table 3.2. Expression trials were conducted for both the His-tagged and un-tagged *Tb*iPGAM, and will be discussed in detail in a latter section (see section 3.6).

Table 3.2 *Tb*iPGAM parameters which were obtained from the Expasy database.

Protein	Molecular mass (Da)	Theoretical pI	Number of amino acids	Extinction coefficient (M^{-1} cm^{-1}) *
His-tagged *Tb*iPGAM	62852.1	5.98	571	39100
Un-tagged *Tb*iPGAM	60603.5	5.56	551	39100

*As measured in water at A_{280} nm

ii) Un-tagged *Tb*iPGAM expression system

The construction of the un-tagged *Tb*iPGAM expression plasmid began with digestion of the original pET28a plasmid containing the *Tb*iPGAM gene with a His-tag engineered at the N-terminus. This was performed by digesting both *Nde*I and *BamH*I sites, which were present in the MCS, in the original plasmid. The resulting gene fragment was subsequently inserted into the pET30a plasmid, which also possesses the corresponding sites (Figure 3.9). The success of the cloning procedure was evaluated by digesting the recombinant clone with the same enzymes: *Nde*I and *BamH*I where DNA fragments corresponding to the size of the vector (5.422 kb) and insert (1.653 kb) were observed (Figure 3.10 a). Moreover, PCR analysis confirmed the existence of the ~0.7 kb amplified product which corresponds to the size of *Tb*iPGAM gene (Figure 3.10 b). The correct sequence was later confirmed by the Dundee Sequencing Service, the University of Dundee. The map of the resulting plasmid is shown in Figure 3.11.

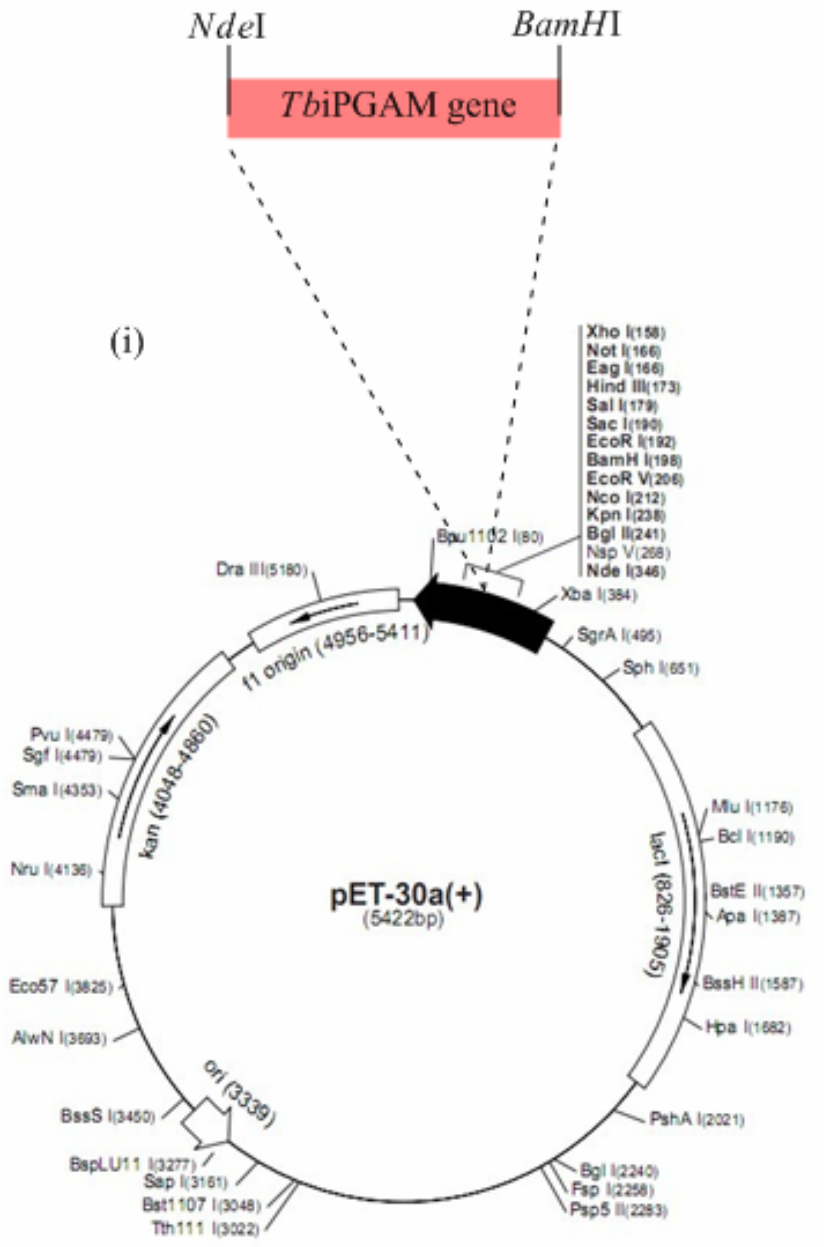


Figure 3.9 The cloning strategy for *Tbi*PGAM which involved the insertion of *Tbi*PGAM gene from the original pET28a plasmid into the final vector (pET30a).

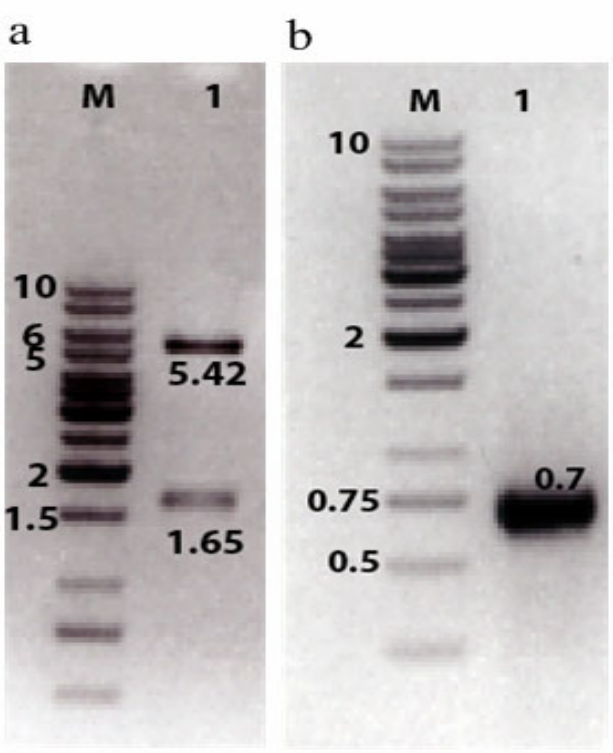


Figure 3.10 Agarose electrophoresis (0.8% (w/v)) of the PCR amplified *Tb*iPGAM and the digestion results. a) a double digested pET30a containing *Tb*iPGAM with *Nde*I and *BamH*I (well 1); and b) PCR amplification of the *Tb*iPGAM gene in the pET30a plasmid. GeneRuler^{TM} 1kb DNA ladder is shown as M.

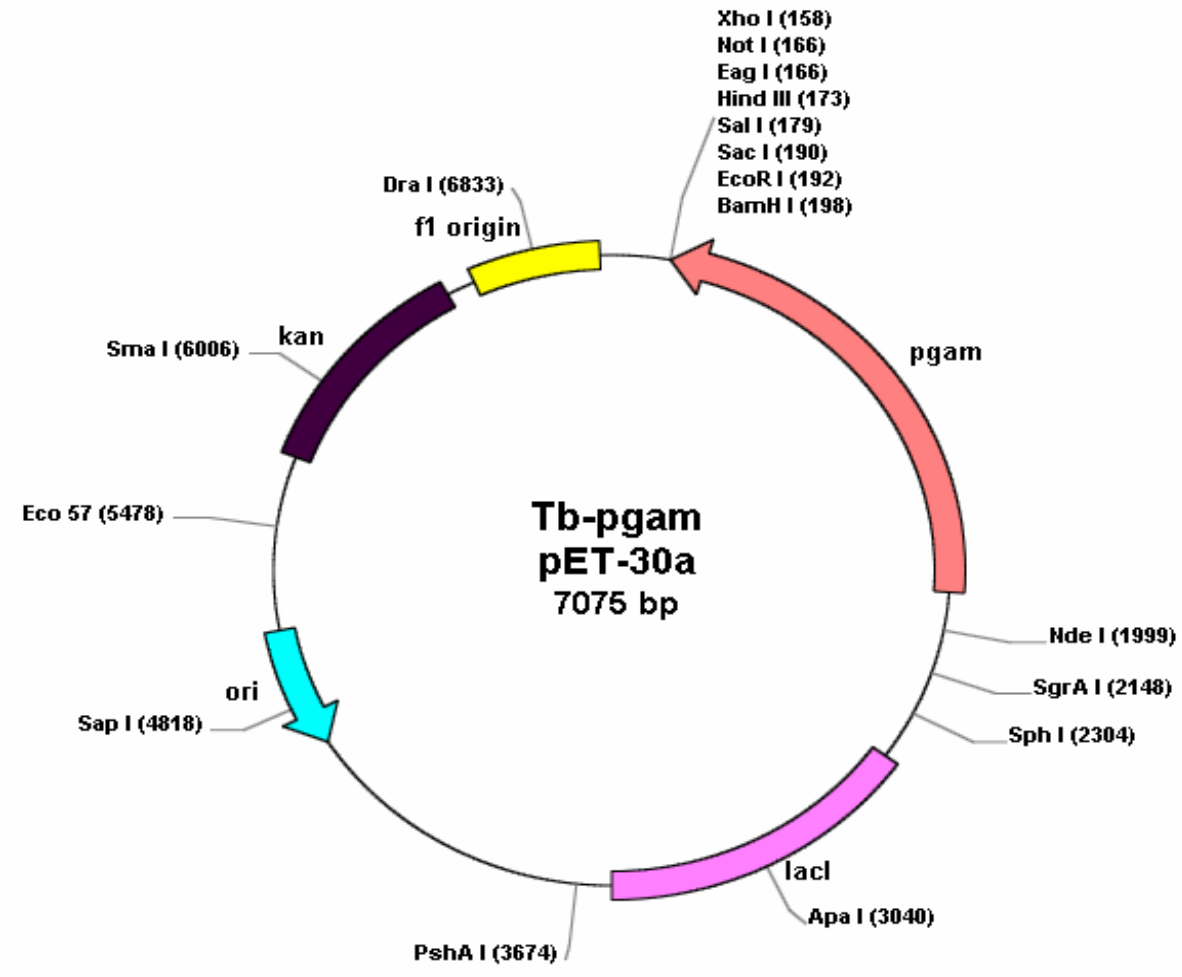


Figure 3.11 The plasmid map of the constructed un-tagged *Tb*iPGAM.

3.6 *Tb*iPGAM expression trials

Various organisms have been manipulated as tools for the production of proteins from plasmids. To date, bacterial (Baneyx 1999), yeast (Cregg et al. 2000; Malys et al. 2011), insect (Kost et al. 2005) and mammalian cells (Rosser et al. 2005; Lackner et al. 2008) have been widely utilised as expression systems for proteins. Among these, *E. coli* is the most extensively used system for the yield of bacterially expressed proteins, as it possesses the advantages of economical value and well-characterised genetics (Ghosh et al. 2004). However, the major downside of the system is that the expressed proteins often occur as insoluble aggregates of folding

intermediates, known as inclusion bodies. It is hence important to determine the optimal expression conditions, in which the solubility of the target protein can be increased. In the laboratory, small-scale expression trials were routinely employed, in order to obtain the most optimal conditions for subsequent large-scale expression and purification.

The expression of soluble *Tb*iPGAM, as reported in Chevalier et al. 2000, was obtained at low production levels, despite various attempts to obtain yields. This study, together with Poonperm 2005 reported the requirement of betaine, which was proposed to stabililise protein folding within cells, and also sorbitol, which plays a functional role in assisting betaine to enter the *E. coli* cells. Nevertheless, it is important to obtain optimal conditions for higher expression levels for this protein, and various expression trials were conducted. In the present study, several expression parameters were included during the expression trials: temperature, cell lines and growth media. In this initial experiment, neither betaine nor sorbitol was included in the primary samples. Different conditions are enumerated in Table 3.3.

Table 3.3 Expression conditions for both the His-tagged and un-tagged *Tb*iPGAM.

Expression condition	**Specification**
Temperature	18°C, 30°C and 37°C
Cell lines	BL21 (DE3), Rosetta 2 (DE3) pLysS, BL21 (DE3) Gold, and BL21 (DE3) CodonPlus
Growth media	LB and 2XYT

Initially, each cell line was included in the expression trials with three distinct temperatures and two types of growth media for the His-tagged and un-tagged *Tb*iPGAM. As a negative control, samples in the absence of IPTG induction were included in each experimental condition. The most optimum condition was determined by SDS-PAGE analysis, where protein bands with the greatest intensity were carefully observed. The small-scale cell lysis procedure was standardised for all samples to ensure uniformity of the results obtained, as was the amount of protein in

each sample to be analysed by SDS-PAGE (see chapter 2). The most favourable conditions for both the His-tagged and un-tagged *Tb*iPGAM in soluble fractions are shown in Figure 3.12, alongside the negative control sample. Prior to this, the insoluble and the soluble fractions were separated by the centrifugation step. The expression conditions found to be favourable for large-scale purification are tabulated in Table 3.4.

Table 3.4 The optimised expression conditions for *Tb*iPGAM.

Expression condition	**Specification**
Cell line	BL21 (DE3)
Temperature	18°C
IPTG concentration	1 mM
Induction time	23 hours
Additional induction regimes	Not required
Shaking speed	170 rpm
Growth media	LB
Harvesting or lysis method	Enzymatic and sonication

The findings showed that the greatest soluble *Tb*-iPGAM expression for both the His-tagged and un-tagged versions were observed at a low temperature (18 °C) with longer incubation period (23 hours), as opposed to higher temperatures with shorter incubation periods. This may be due to the increment in cell viability at lower temperature, as well as the enhancement of the protein folding pathway (Li et al. 2001). The relative expression of the un-tagged *Tb*iPGAM was better than the His-tagged version in BL21 (DE3) cells, although in other cells lines (BL21 (DE3) Gold and BL21 (DE3) CodonPlus), both versions exhibited similar expression level. Nevertheless, the lowest expression level was observed in Rosetta 2 (DE3) pLysS, for both the un-tagged and His-tagged versions. It is noteworthy that a sonication step was necessary, as an additional cell lysis procedure alongside lysozyme treatment. The use of shaker flasks instead of Falcon tubes and a slower shaking

speed also increased the solubility of the proteins. Higher expression levels in the soluble fractions for both the His-tagged and un-tagged *Tb*iPGAM were successfully obtained, increasing the chances for obtaining higher protein yields during purification.

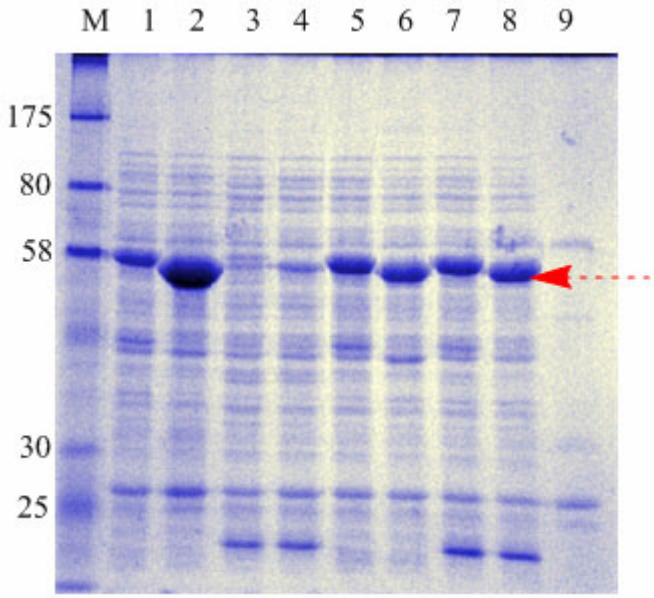


Figure 3.12 *Tb*iPGAM expression trials at 18°C for 23 hours. A 12% SDS-PAGE is shown with 2-4 µg samples of IPTG-induced samples containing *Tb*iPGAM. Lanes 1 (His-tagged) and 2 (un-tagged) are BL21 (DE3), lanes 3 (His-tagged) and 4 (un-tagged) are Rosetta 2 (DE3) pLysS, lanes 5 (His-tagged) and 6 (un-tagged) are BL21 (DE3) Gold and lanes 7 (His-tagged) and 8 (un-tagged) are BL21 (DE3) Codon Plus. The red arrow indicates the over expressed *Tb*iPGAM.

3.7 Purification steps for *Tb*iPGAM

The high percentage of sequence identity between *Lm*iPGAM and *Tb*iPGAM suggested that both proteins may exhibit similar behaviour during purification stages. Hence, it was decided that a similar approach should be employed for the latter enzyme. In this experiment however, *Tb*iPGAM eluted from IMAC was divided into two separate additional purification steps: i) desalting and ion-exchange steps analogous to *Lm*iPGAM purification or ii) an additional gel filtration step without the other two polishing steps.

i) Step one: immobillised metal-affinity chromatography

As discussed previously in section 3.4 i, the cell lysate containing *Tb*iPGAM was applied to a 5 ml Fast Flow IMAC column (GE Healthcare). As intended, the His-tagged *Tb*iPGAM attached to the matrix, while the unbound proteins were washed away in the flow-through in the absence of imidazole. The increase of imidazole concentration to 25 mM resulted in weakly-bound proteins being eluted, as shown by the small UV peak at A_{280} nm (Figure 3.13 a). A step-wise increase to 500 mM imidazole eluted a sharp UV peak at A_{280} nm corresponding to essentially pure *Tb*iPGAM (Figure 3.13 b). The total protein yield from 1 litre of culture was determined to be 59 mg.

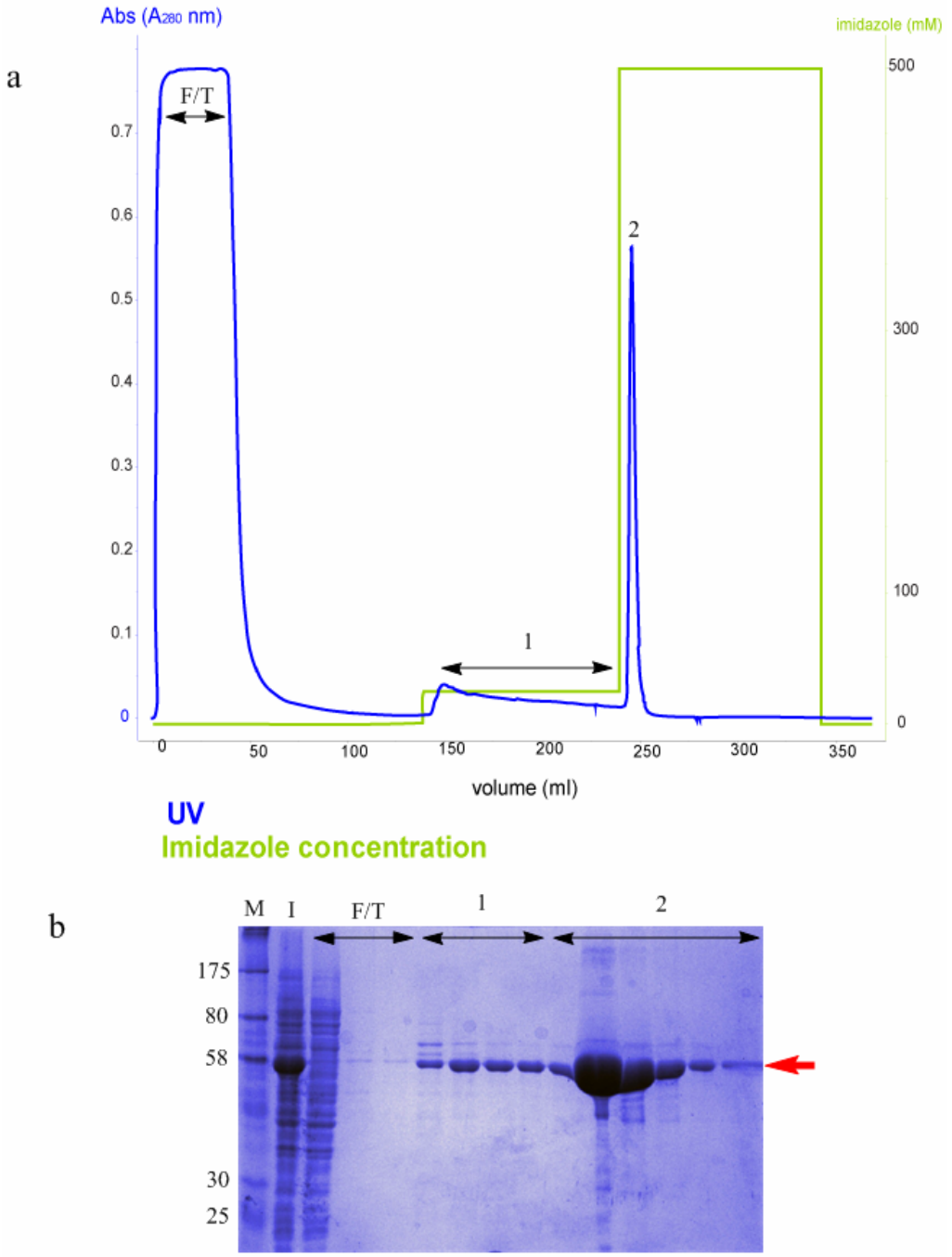


Figure 3.13 a) The elution profile of *Tb*iPGAM from a 5 ml Fast Flow IMAC column resulting in two distinct elution sections (1 and 2) as well as an earlier flow-through (F/T) fractions (blue trace). b) A 12% SDS-PAGE is shown with 2-4 µg samples of each protein fraction with initial extracts (I), flow-through (F/T) and eluates from sections 1 and 2. The red arrow shows protein corresponding to the size of His-tagged *Tb*iPGAM (62.9 kDa). The protein markers are shown as M and labelled as kDa.

ii) Step two: desalting

Pooled fractions of *Tb*iPGAM from the IMAC step were applied to a HiPrep 26/10 column, a pre-packed Sephadex G-25 Fine column for desalting. As depicted in Figure 3.14 a, a large UV peak at A_{280} nm corresponding to *Tb*iPGAM was observed. The subsequent increase in conductivity (brown trace) indicates the elution of salts or any other smaller molecules, as also observed in Figure 3.7 for *Lm*iPGAM. Through SDS-PAGE analysis (Figure 3.14 b) a conspicuous band corresponding to the size of *Tb*iPGAM was observed, indicating the presence of essentially protein suitable for the subsequent chromatography step.

iii) Step three: ion-exchange chromatography

An identical procedure as *Lm*iPGAM was conducted for *Tb*iPGAM, where a shallow gradient elution was performed by increasing the concentration of NaCl to a maximum of 1 M. The resulting UV protein peaks are illustrated in Figure 3.15 a, where similar elution pattern to that for *Lm*iPGAM was obtained, albeit with poorer separation and different relative peak heights. The peak differences from the previous *Lm*iPGAM experiment may indicate that both enzymes experienced different distribution of conformers or oligomers. Nevertheless, it is also plausible that the length of elution has to be increased, to obtain a better peak separation, as the resolution of the column is identical to the previous ion-exchange chromatography for *Lm*iPGAM. SDS-PAGE analysis exhibited essentially identical protein bands corresponding to the *Tb*iPGAM molecular mass in all the fractions (Figure 3.15 b). Lower molecular mass bands were also observed, which may well be the consequences of degradation.

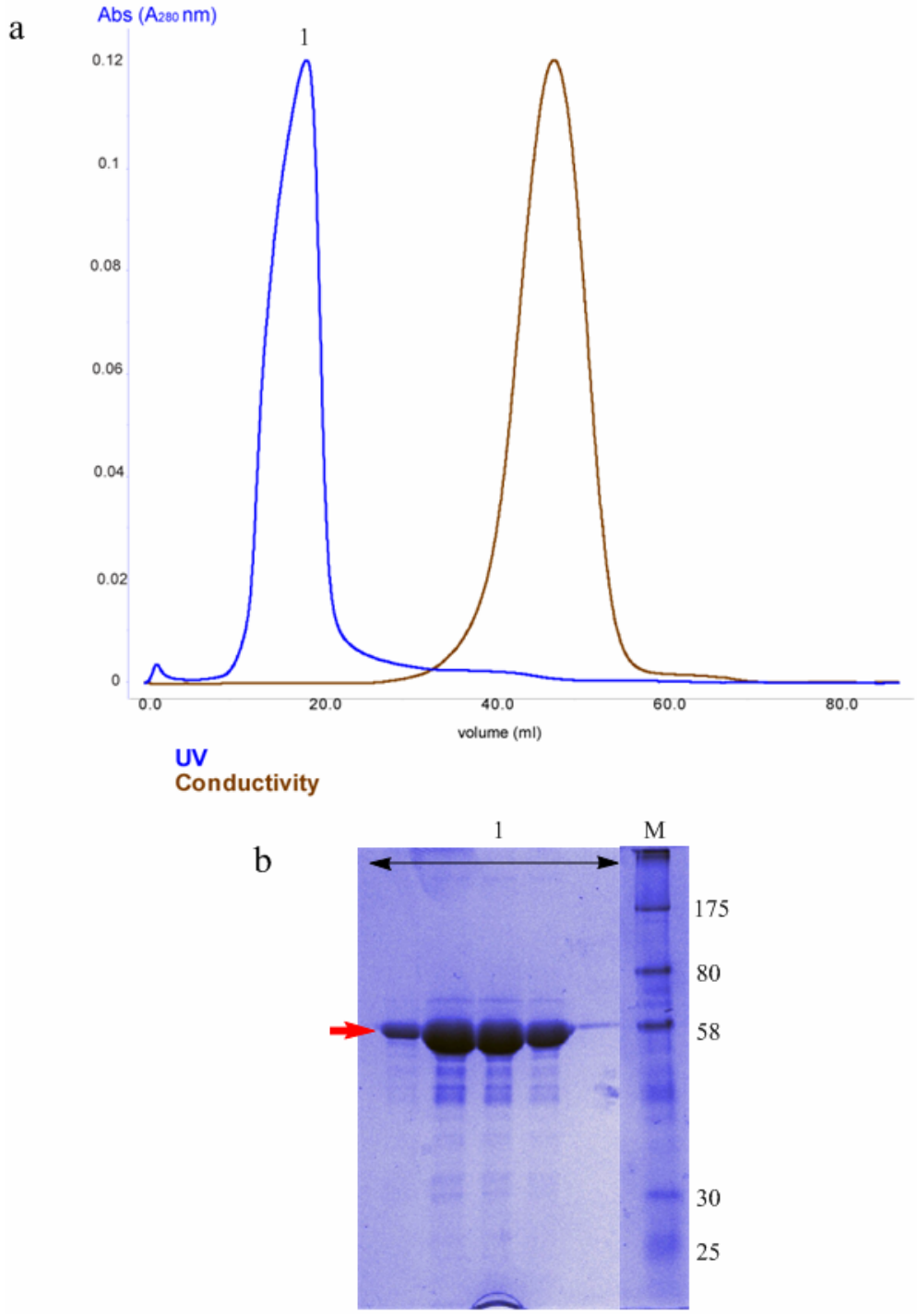


Figure 3.14 a)The elution profile of *Tbi*PGAM from the HiPrep 26/10 desalting column showed a single protein peak (blue trace), and a peak with salt and smaller molecules (brown trace). b) A 12% SDS-PAGE is shown with 2-4 µg samples of each protein fraction corresponding to the blue trace. The red arrow indicates the proteins corresponding to the size of His-tagged *Tbi*PGAM (62.9 kDa). The protein marker is shown as M and labelled as kDa.

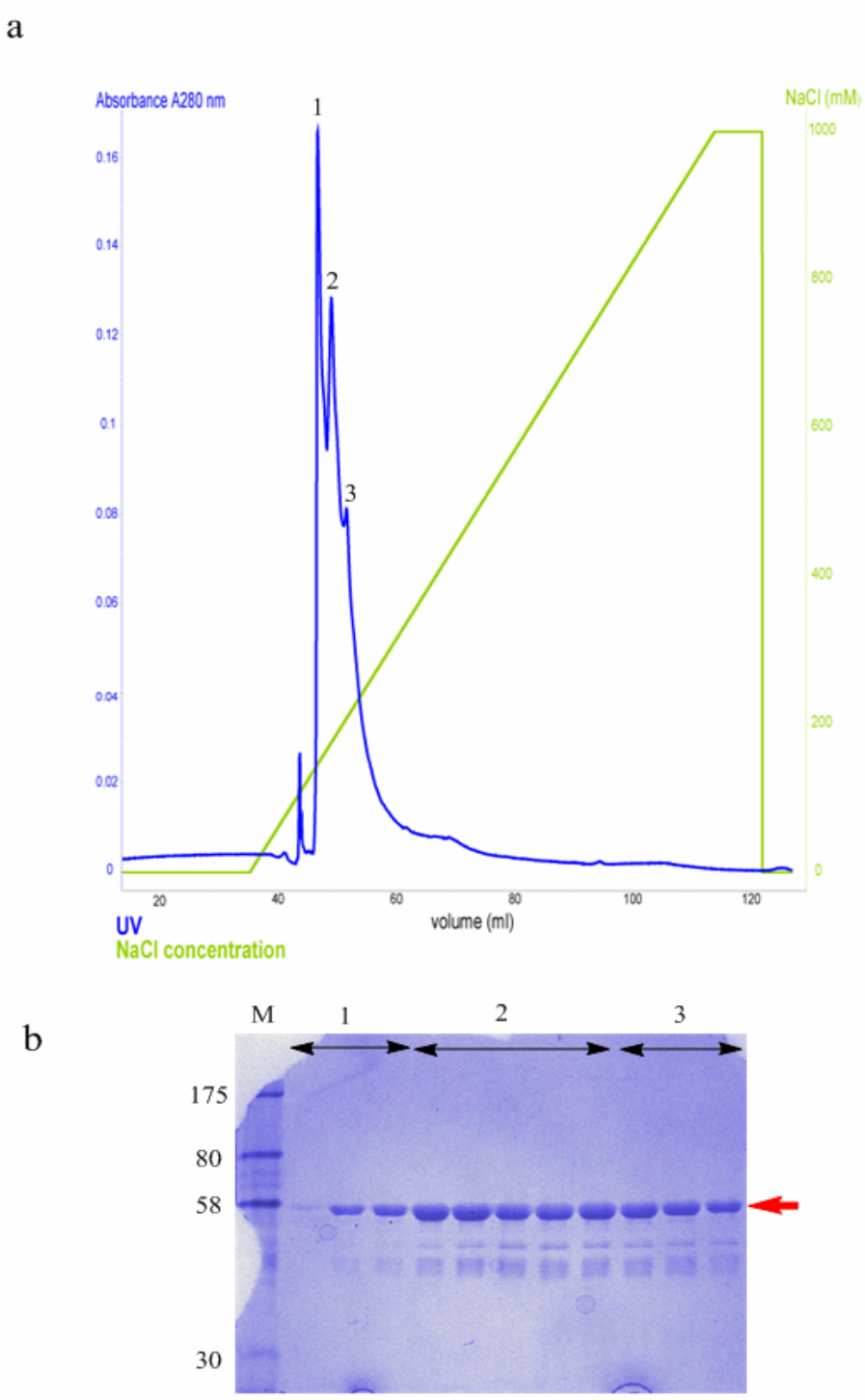


Figure 3.15 a) The elution profile of *Tb*iPGAM from the MonoQ 5/50 GL column showed three poorly-separated peaks (blue trace), labelled 1, 2 and 3. b) A 12% SDS-PAGE is shown with 2-4 µg samples of each protein fraction (blue trace). The red arrow shows proteins corresponding to the size of His-tagged *Tb*iPGAM (62.9 kDa). The protein markers are shown as M and labelled as kDa.

iv) Alternative to desalting and ion-exchange steps: Gel filtration

A portion of essentially pure *Tb*iPGAM from the initial IMAC purification was applied to a Superdex 200 10/300 GL column, for post-purification analysis as well as for a 'polishing' step. As depicted in Figure 3.16 a, two UV peaks could be observed, although it is apparent that peak separation was not optimal. Samples corresponding to the eluted fractions were analysed individually by SDS-PAGE (Figure 3.16 b), and indicated pure protein bands conforming to the molecular mass of *Tb*iPGAM. Kinetic measurements of *Tb*iPGAM were carried out using pooled fractions from the major peak 2, and will be discussed in chapter 6.

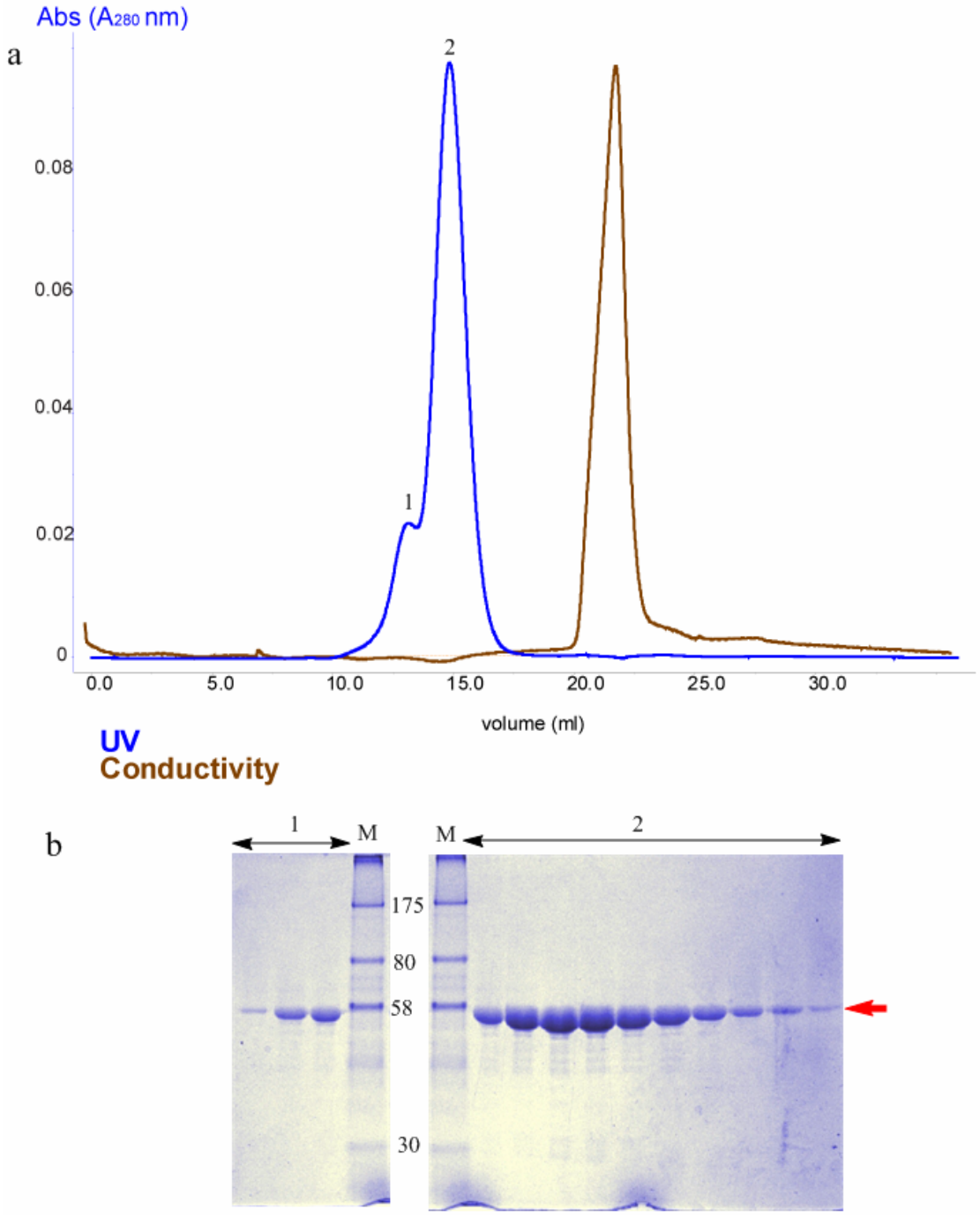


Figure 3.16 The elution profile of *Tbi*PGAM from a Superdex 200 10/300 GL column showed two overlapping peaks (blue trace) labelled 1 and 2, as well as a peak containing salt and smaller molecules (brown trace). b) A 12% SDS-PAGE is shown with 2-4 µg samples of each protein fraction corresponding to the blue trace. The red arrow shows the pure proteins corresponding to the size of His-tagged *Tbi*PGAM (62.9 kDa). The protein marker is shown as M and labelled as kDa.

3.8 Conclusion

As the primary step in any protein characterisation analysis, purification methods required careful optimisation in order to yield high purity proteins. The chromatography method for macromolecular separation has to be wisely chosen, in order to obtain pure and well-separated macromolecules, either in the initial or polishing step of purification. An IMAC column provides an excellent tool to obtain the initial specific protein of interest, but nevertheless required post-purification steps to eliminate similar proteins and to exchange into the correct buffer.

Both His-tagged and un-tagged *Lm*iPGAM and *Tb*iPGAM expression systems have been developed by the Brussels and Edinburgh research groups, respectively. Only the *Lm*iPGAM His-tagged proteins were chosen to be purified, because the method was more efficient in yielding high-purity protein samples. Despite that the un-tagged *Tb*iPGAM overexpression was better than the His-tagged version, the same purification approach as *Lm*iPGAM was adopted, due to the efficiency of the protocol and to provide direct comparisons with the former enzyme. Ni^{2+} was selected as the metal ion to interact with the His tags, as opposed to Co^{2+} in the previous single-step purification protocol implemented by Chevalier et al. 2000, Guerra et al. 2004 and Poonperm 2005. This modification is vital to avoid any residual Co^{2+} ions to be entrapped in the pure proteins. Subsequently, a desalting step was employed to remove imidazole and excess salt, followed by an ion-exchange chromatography. SDS-PAGE analysis revealed the presence of high purity *Lm*iPGAM and *Tb*iPGAM at each purification stage, indicating the probability of obtaining decent structural and biophysical data for both trypanosomatid iPGAMs. It is noteworthy that the purity and yield of *Lm*iPGAM which were obtained in this study was similar to the previous literature (Poonperm 2005), but the additional ion-exchange chromatography step provides the possibility of achieving different conformations or oligomeric states of the same protein.

It is indeed interesting to note that ion-exchange chromatography revealed the possibility for *Lm*iPGAM to be further separated based on its biochemical and biophysical properties, as will be described in chapter 4. The yield of pure

*Lm*iPGAM was found to be 20 mg from one litre of cell culture following the standard expression method, whereas the *Tb*iPGAM yield was significantly higher with 59 mg from a litre of cell culture. The higher *Tb*iPGAM production compared to the previous literature (Chevalier et al. 2000; Poonperm 2005) may well be the consequence of the improved expression conditions achieved in this current study.

Chapter 4:

The Different Forms and Multiple Oligomeric States of Trypanosomatid iPGAMS

4.1 Aims

The cofactor-independent iPGAM from *L. mexicana* has been characterised extensively, with many observations being related to its dependency for Co^{2+}, as well as to structural analyses of the existing closed conformation (Poonperm 2005; Guerra et al. 2004; Nowicki et al. 2009). These findings have improved an understanding of the structure, function and mechanism of action for the enzyme tremendously, thereby providing significant encouragement for trypanosomatid iPGAM as a drug target. Nevertheless, from a structural point of view and in relation to drug design, knowledge of the open-form structure would be equally beneficial for virtual screening. This is particularly true for *Lm*iPGAM because the active site in the closed conformation in the presence of substrate and Co^{2+} is very small, and is poorly accessed from the solvent, while the open conformation possesses a wider cleft between the two domains (Figure 4.1). It is vital therefore to characterise the behaviour of the enzyme biophysically, in order to identify conditions in solution in which *Lm*iPGAM adopts the open conformation for the purpose of drug design and screening.

The previous chapter has discussed the three significant well-separated fractions eluted from the anion-exchange column, as post purification to Ni^{2+}-IMAC and desalting steps. The iPGAM samples in P1, P2 and P3 were indistinguishable by SDS-PAGE, so further analysis was required to identify the reasons for the different behaviour on the ion-exchange column. The differences in the elution volume may well be correlated to the metal binding and content of the enzymes. Apart from this, it was also plausible that the differences between each peak were the consequence of different oligomeric states or variations in protein folding, where this eventually leads to different forms of *Lm*iPGAM. Hence, the main aim for this chapter was to characterise the three distinct peaks from the ion-exchange column, which may provide some new insights into trypanosomatid iPGAM behaviour.

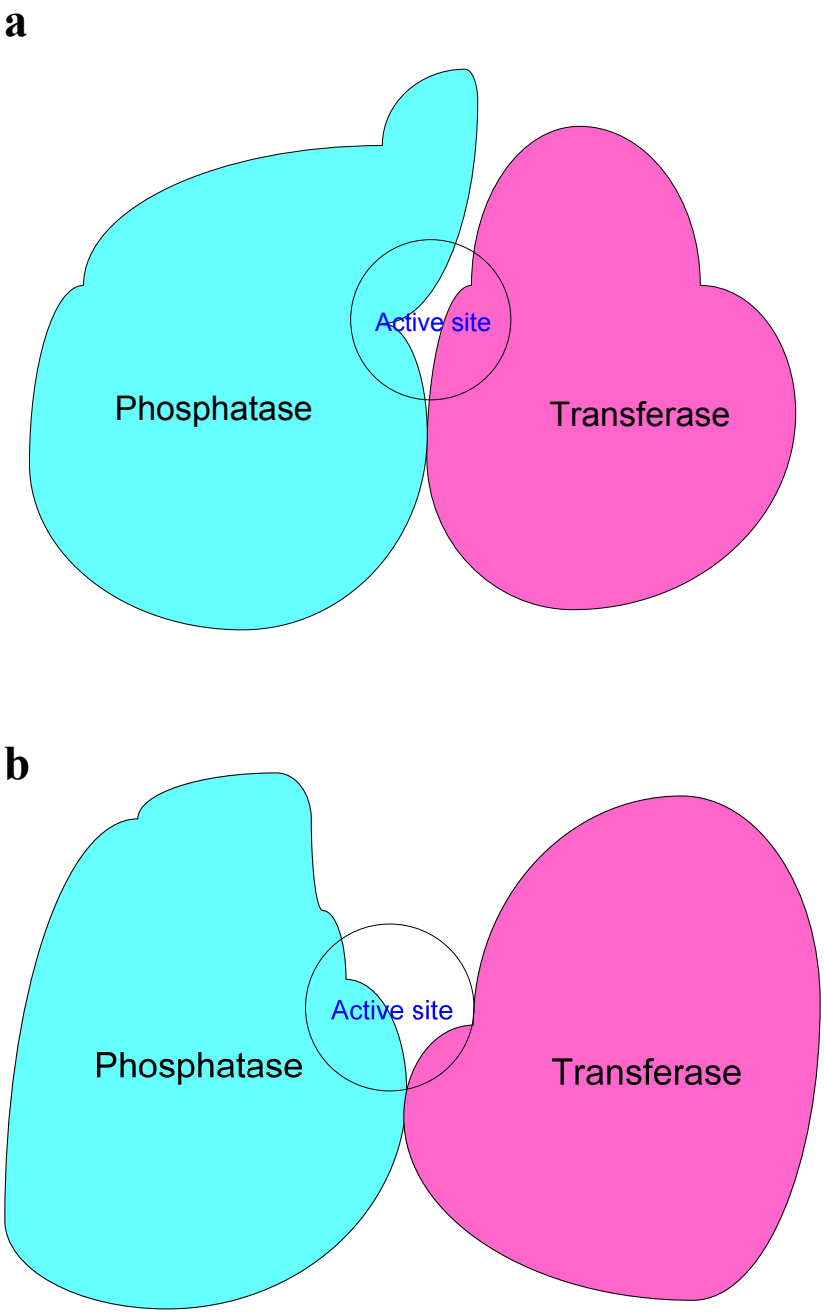


Figure 4.1 a) The closed form iPGAM with a poorly accessible active site
b) The open form iPGAM with a more accessible active site.

4.2 Introduction

Proteins start to fold during the translational process, and may experience modifications before, during, or after release from the ribosomes. It has been proposed that protein chains, or polypeptides, obtain three-dimensional structure as well as altered dynamics during the translational process (Ellis et al. 2009; Gomes and Wittung-Stafshede 2011). From the ribosome, the majority of proteins fold into their characteristic functional three-dimensional structures with their preferred functional shapes or conformations (Alberts et al. 2008). This process is extremely important since the three-dimensional structure is essential for a protein's functional role. Nevertheless, conformational transitions can occur and this normally involves the relative movement of almost rigid structural elements. In the majority of proteins, these domain motions are explicitly vital for protein functions, specifically for enzymes in catalysis (Tama and Sanejouand 2001). An example of such phenomena has been observed in citrate synthase, which is a two-domain protein. The binding of coenzyme A induces an 18° rotation of the small domain around an axis close to residue 274 that represents the hinge (Remington et al.1982; Huber and Bennett 1983; Wiegand and Remington 1986; Tama and Sanejouand 2001). Consequently, this triggers the closure of the cleft between the two domains, where the substrate binding site lies (Tama and Sanejouand 2001). Moreover, mutations of specific residues resulting in amino acid substitutions can also alter the equilibrium between the open and closed conformations of some proteins (Nallamsetty and Waugh 2007). Often too, these conformational transitions consequently effect the activation and inactivation of protein, as observed in some cases (Wu et al. 2009; Olsson and Wolf-Watz 2010).

Proteins have also evolved through oligomerisation phenomena, where two or more polypeptide chains were associated together (Ali and Imperiali 2005). Various factors contributing to dimerisation, as well as the effects on the structure and function of proteins have been extensively discussed (Marianayagam et al. 2004; Ali and Imperiali 2005). This is mostly related to molecular stabilisation, supporting functionality, and sometimes reported to be disadvantageous to the enzyme activity. In this chapter, further characterisation of iPGAM will unravel the potential of the

previously reported monomeric enzyme to form a dimer, which is a new development in understanding the behaviour of iPGAM.

4.3 Modelling the *L. mexicana* iPGAM open-form structure

In the absence of the availability of a crystal structure of the open form of trypanosomatid iPGAM, a model had to be generated. This was important to provide the basis for molecular interpretations of experimental data. Previously, the closed-form structure of *L. mexicana* iPGAM was solved in the presence of the substrate 3PGA at both high and low Co^{2+} concentrations (Poonperm 2005; Nowicki et al. 2009). Nevertheless, as mentioned above, an open-form model of *Lm*iPGAM would also be important, to facilitate a preliminary understanding of the structural variations of the enzyme, as an outcome of domain movements. Prior to the work described in this thesis, the open-form crystal structure of iPGAM from *B. anthracis* (Nukui et al. 2007) had been obtained, and was the best option to generate an *Lm*iPGAM model. As described in chapter 1, *Ba*iPGAM has ~33% sequence identity to *Lm*iPGAM, and was the closest open-form structure available for the modelling at the time. However, more recently, an open-form structure of *T. brucei* iPGAM (PDB ID 3NVL) (Mercaldi et al. 2012) with its ~78% identity to *Lm*iPGAM has been reported, which has shed further light on understanding the almost rigid yet flexible domain movements of the enzyme. The next section will focus on the modelling of open-form *Lm*iPGAM based on the crystal structures of *Ba*iPGAM and *Tb*iPGAM.

i) Modelling based on *B. anthracis* open-form structure

The crystal structure of the open form of *Ba*iPGAM comprises two distinct domains: phosphatase and transferase domains, connected by two short peptide linkers: linker 1 and 2 (Figure 1.5 chapter 1). The crystal structure was solved at 2.40 Å and exhibits the occupation of Mn^{2+} in both metal sites in the presence of a water molecule. This structure was shown to be in the open conformation when compared to the previous *B. stearothermophilus* iPGAM structure which was solved in the presence of 3PGA (Jedrzejas et al. 2000a) or 2PGA (Jedrzejas et al. 2000b), with two Mn^{2+} ions in each structure.

The open-form *Ba*iPGAM structure was used as a template to predict the open-form model of *Lm*iPGAM by using Friend for MODELLER, a program used for homology and comparative modelling of protein three-dimensional structures by using an integrated analytical front-end application for bioinformatics (Abyzov et al. 2005). The method used to generate the structure has been described in chapter 2. Figure 4.2 illustrates the superposition of the model of *Lm*iPGAM with the closed form crystal structure (PDB ID 3IGY). The recent publication of the open-form structure of *Tb*iPGAM has shown that its structure is substantially different from that of open-form *Ba*iPGAM, and it is likely that the more closely related *Tb*iPGAM now provides a better model for *Lm*iPGAM (see next section).

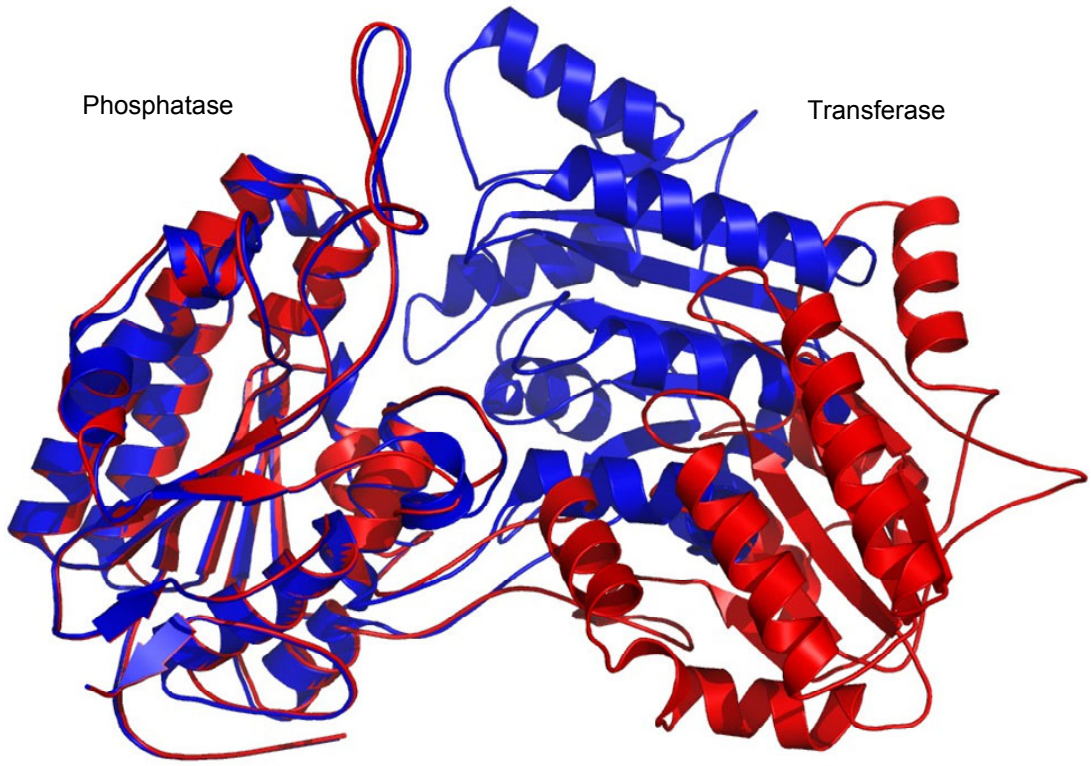


Figure 4.2 The overlay of the closed-form structure of *Lm*iPGAM (PDB ID 3IGY) (blue) with the generated open-form model (red) based on *Ba*iPGAM (PDB ID 2IFY).

ii) *T. brucei* iPGAM open-form structure

It is indeed interesting for *Lm*iPGAM to be compared to the recently reported open-form monomeric structure of *Tb*iPGAM. It is likely that both enzymes can exist in similar forms, either in the closed or open states, since both shares a high percentage of sequence identity (78%). As depicted in Figure 4.3, the structure of *Tb*iPGAM consists of the two typical phosphatase and transferase domains of iPGAMS. The prediction of the hydrodynamic radius of the open form of *Lm*iPGAM will be more accurately based on the *Tb*iPGAM structure, rather than on that of *Ba*iPGAM.

The structural comparisons suggest the possibility that trypanosomatid iPGAMs (and possibly plant iPGAMs) belong to one family, while the bacterial iPGAMs belong to another different family. A sequence alignment of iPGAMs from various organisms (see chapter 1) indicates that the enzyme could be divided into two distinct families, where trypanosomatid iPGAMs are grouped together with plant iPGAMs, whereas the iPGAMs from bacteria, nematodes, fungi and algae could be categorised in another family (Linda Gilmore, unpublished data). This suggestion coincidentally fits with the proposed one-metal mechanism (Nowicki et al. 2009), which suggests that iPGAMs from trypanosomatids utilise only one metal site during catalysis, which is suspected to be similar with plant iPGAMs. The bacterial iPGAMs (Jedrzejas et al.2000a; Jedrzejas et al. 2000b) on the hand, function with a two-metal mechanism. It is plausible that the bacterial enzymes possess a larger open conformation because the two metal sites are involved in catalytic function, whereas trypanosomatid iPGAMs perform their function with a further reduced open conformation, with a proposed one-metal mechanism during catalysis.

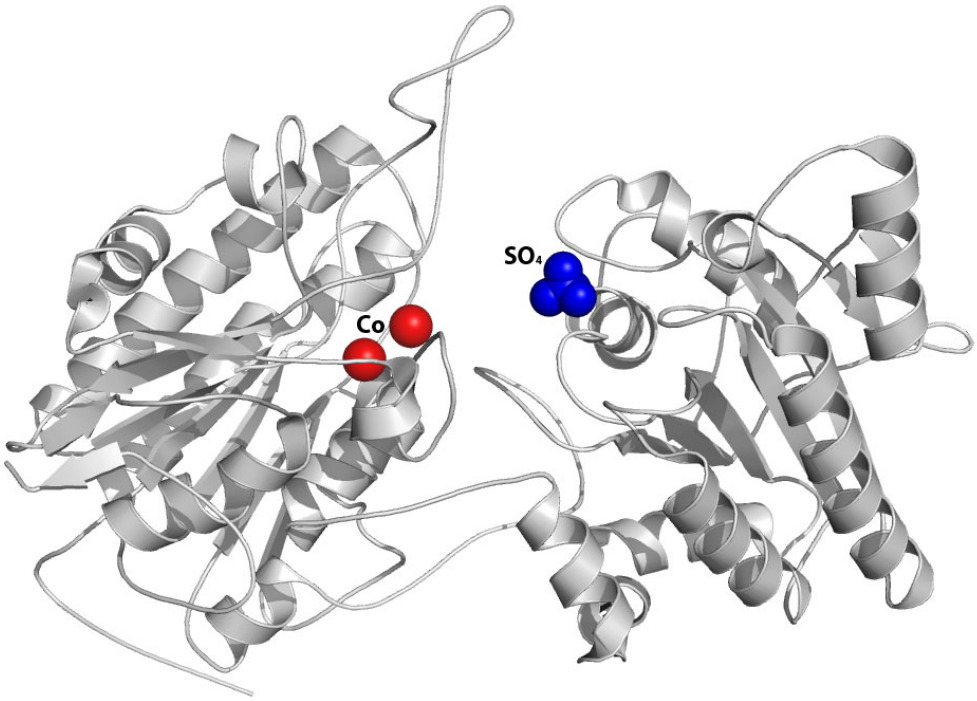


Figure 4.3 *Tb*iPGAM (PDB ID 3NVL) crystallised in its open conformation (Mercaldi et al. 2012).

4.4 Hydrodynamic radii of iPGAMs

i) Closed-form crystal structure of *Lm*iPGAM and open-form crystal structure of *Tb*iPGAM

The hydrodynamic size of the open and closed conformation, as well as the monomer and dimer states of the enzyme can be predicted using the solved crystal structures. This can be performed by using HYDROPRO software (Garcia de la Torre et al. 2000; Ortega and Garcia de la Torre 2007; Ortega et al. 2011), which enables the measurement to be employed as an indication of the hydrodynamic properties of the globular proteins from their atomic level structure. The calculated measurement however, relies on a number of geometric assumptions, such as shapes, which can be either simple (ellipsoids and cylinders) or more arbitrarily complex (Ortega and

Garcia de la Torre 2007), and may deviate from the experimental values. On the other hand, PyMol software (DeLano Scientific 2002) can be used to determine the maximum dimension of proteins in the crystal lattice. It is noteworthy that the maximum dimensions from the crystal structures of monomeric *Lm*iPGAM, *Tb*iPGAM open-form monomeric structure, and *Ba*iPGAM open-form monomeric structure using the latter method were performed without standardising specific residues in each molecule. Nevertheless, this was done by selecting the widest distance that corresponds to the rotational diameter of the proteins.

Table 4.1 The hydrodynamic properties for the closed-form monomeric *Lm*iPGAM, and the open forms of both *Tb*iPGAM and *Ba*iPGAM. The values were obtained from the HYDROPRO software, and from PyMol representations of the crystal structures. The molecular masses used in this analysis correspond to the un-tagged version of the proteins.

Protein	**PDB**	**Molecular mass (Da)**	**Radius of gyration (R_g)(nm)**	**Stoke's radius (R_h)(nm)**	**Volume (nm^3)**	**Maximum distance(nm)**
*Lm*iPGAM	3IGY	60723	2.49	3.21	95.84	8.81
*Tb*iPGAM	3NVL	60604	2.64	3.41	98.72	9.11
*Ba*iPGAM	2IFY	56149	2.79	3.61	93.86	9.58

The results from HYDROPRO show that the open-form *Ba*iPGAM is the largest molecule, with Stoke's radius (also known as hydrodynamic radius, R_h) of 3.61 nm, followed by the open-form *Tb*iPGAM (3.41 nm) and the closed-form *Lm*iPGAM (3.21 nm). The radius of gyration (R_g) from the three samples also display similar pattern in all samples, where *Ba*iPGAM >*Tb*iPGAM > *Lm*iPGAM (Table 4.1). In this case, the apparent differences between the two types of radii are due to the consideration of the hydration layer for hydrodynamic radius, which includes the effects of both the solvent (hydro) and shapes (dynamic) (http://ecoserver.imbb.forth.gr/pdf/hydrodynamic_radius.pdf). The radius of gyration, on the other hand, is a measurement of the root mean square radius of the molecule from a given axis or its centre of gravity. Both radii are illustrated in Figure

4.4. The relationship between R_h and R_g has been discussed by Zif et al. 2008, where a simple geometrical calculation exhibits that for a spherical globule, $R_g/R_h = \sqrt{(3/5)}$. As mentioned above however, there may be some deviations from the experimental results, as also been discussed by Zif et al. 2008. The experimental diffusion measurements of globular proteins exhibit larger hydrodynamic radii than the calculated values, which is a consequence of the solvation shell (a layer of weakly-bound water molecules) that is present around folded proteins (Zif et al. 2008). These comparisons are found to be similar to iPGAM in the later section (see section 4.5 iii).

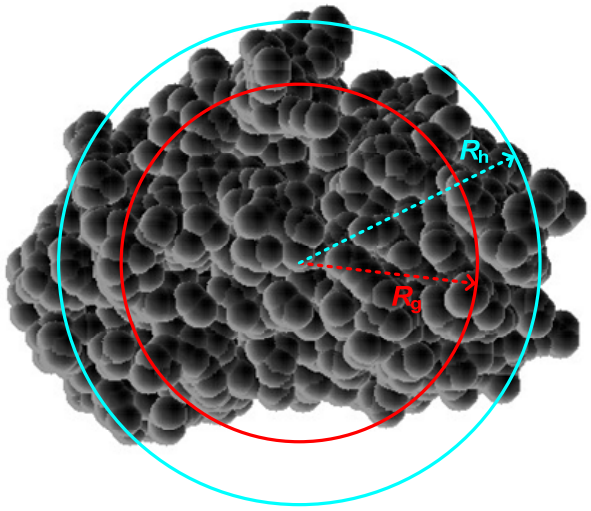


Figure 4.4 The radius of gyration (R_g) and the hydrodynamic radius (R_h) for a spherical globule indicated in red and cyan lines. The protein molecule is the closed-form *Lm*iPGAM.

A significant observation from the structural comparisons shown in Figure 4.5 is that the trypanosomatid iPGAM open conformation (Figure 4.5c) is strikingly different in shape from the bacterial open form (Figure 4.5b). The open conformation of *Lm*iPGAM is predicted to have a very similar hydrodynamic size to that of the open conformation of *Tb*iPGAM because of the strong sequence conservation. *Ba*iPGAM possesses a more "open" conformation than the *Tb*iPGAM open conformation, or the proposed open form of *Lm*iPGAM. On the other hand, the

maximum dimension as measured from the crystal lattice for *Ba*iPGAM is ~8.49 nm. The open-form structure of *Tb*iPGAM exhibited a rotational diameter of ~8.13 nm, followed by the closed form structure of *Lm*iPGAM with ~7.66 nm. Nevertheless, the trends of both the calculated hydrodynamic sizes, as well as the maximum dimensions from the crystal lattice were very similar, where the open conformation of *Ba*iPGAM is larger than that of *Tb*iPGAM, and the closed-form *Lm*iPGAM is the smallest molecule.

Table 4.2 The differences between the calculated hydrodynamic diameter and the maximum dimension from the crystal lattice.

Protein	**Hydrodynamic diameter from HYDROPRO (nm)**	**Maximum dimension from the crystal structure (nm)**
*Lm*iPGAM	6.42	7.66
*Tb*iPGAM	6.82	8.13
*Ba*iPGAM	7.22	8.49

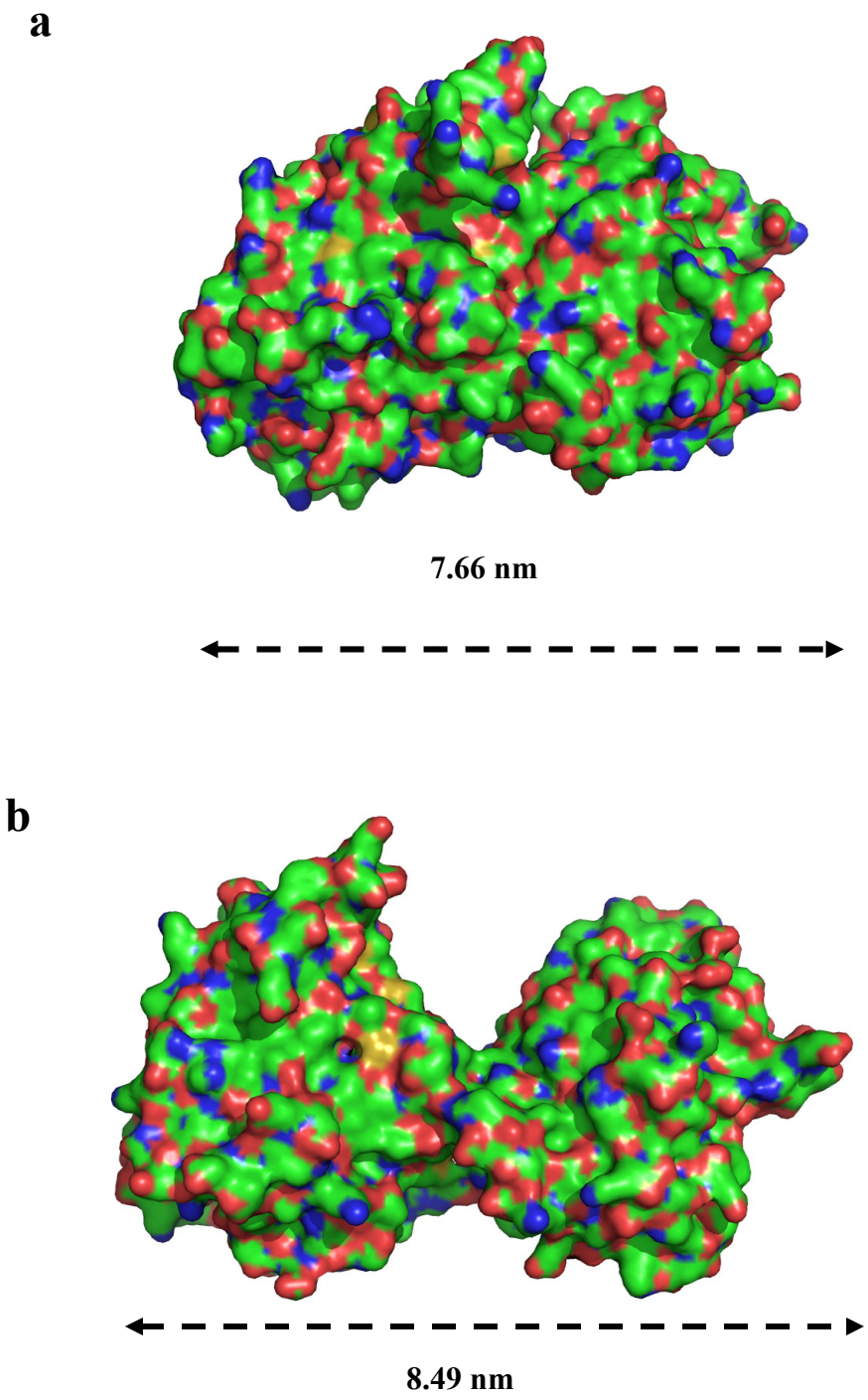


Figure 4.5 The maximum dimensions of a) *Lmi*PGAM (PDB ID 3IGY) monomeric closed-form structure, b) *Bai*PGAM (PDB ID 2IFY) monomeric open-form structure, and c) *Tbi*PGAM (PDB ID 3NVL) monomeric open-form structure

c

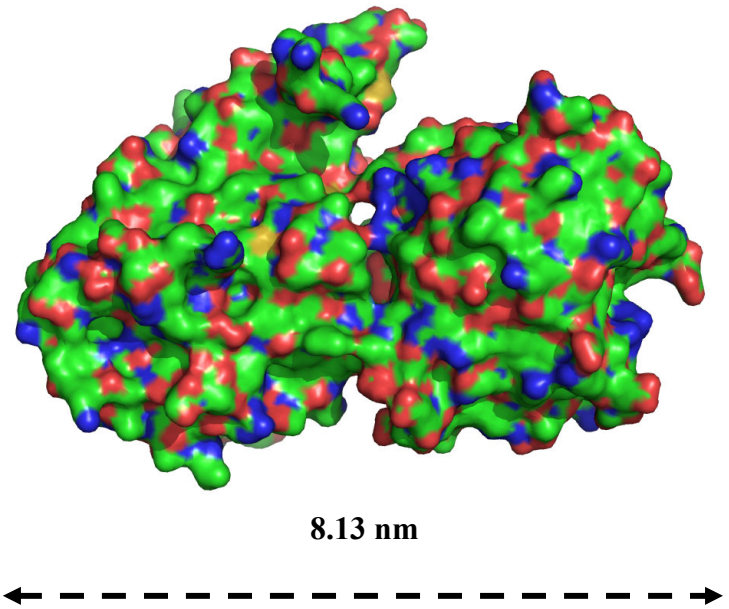


Figure 4.5 contd.

4.5 Characterisation of the peaks from ion-exchange chromatography of *Lm*iPGAM

As described in chapter 3, three well-separated *Lm*iPGAM protein peaks were obtained from the third-step of purification using the anion-exchange column. It is relevant to characterise these peaks using various biochemical and biophysical methods, with a view to identify the reasons for their different elution behaviour. The differences in the elution positions may indicate different charge distribution on the surface of the protein molecules because of distinct structural forms or oligomeric states. These observations may be a result of post-translational modifications of the proteins to form oligomers (Wang et al. 2008), or may be a consequence of conformational transitions of the correctly folded polypeptides to create different forms of the same protein. Nevertheless, all three samples exhibited similar

molecular masses as indicated by SDS-PAGE analysis (see chapter 3). The main question would be whether *Lm*iPGAM, previously reported as a monomer of ~60.7 kDa, has the ability to form oligomers in solution. Hence, in the next sections, the biophysical properties of each peak will be discussed in detail.

i) Specific activity measurement

It was important to ensure that all peaks were active and did not exhibit a loss of activity, which may well be a result of incorrectly folded protein. In this initial experiment, a continuous enzyme assay was employed, in which the activity of *Lm*iPGAM was monitored by observing the decrease absorbance of NADH at A_{340} nm. In this particular experiment, the three samples, P1, P2 and P3 were incubated with 0.1 mM Co^{2+} at 25°C, and the activity measurement was carried out at various time of incubation. The initial (time 0) specific activities for all three samples were observed to be similar: 20, 19.8 and 15.9 $\mu mol.min^{-1}.mg^{-1}$ for P1, P2 and P3, respectively. The activities were also measured after 15 min, 1, 3, 6, 9, 18, 24 and 48 h of incubation.

The results, as shown in Figure 4.6, provide an interesting observation, where the activities of each individual *Lm*iPGAM increased hyperbolically, and displayed significant differences in the trend of activity increment. Although the initial activities of the three samples were similar, P1 exhibited a substantial increase throughout the incubation period, resulting in almost 2-fold to 3-fold higher activity compared to P2 and P3. After 48 hours, the activity of P1 has reached a "plateau" or stationary phase, where the maximum activity of Ni^{2+}-IMAC purified enzyme has almost been obtained. Both P2 and P3 on the other hand, still required longer incubation to reach their maximal activities. These observations indicate a distinctive regulation of activity for P1, compared to P2 and P3. The variation may well be a consequence of the flexibility of the protein around the active site (Fieulaine et al. 2011), where the catalytic site of *Lm*iPGAM in this case, is located in the cleft between two functional domains. This may have resulted in the active and inactive states of the enzyme, which coincidentally could be distinguished using the ion-exchange step. The active state may be correlated to the closed conformation of the

enzyme, which may exist in the presence of the substrate 3PGA, as well as divalent metals which occupy the metal sites. The effect of the substrate and metal will be discussed in section 4.5 iii. The open-form conformation, on the other hand, may well be the inactive state of the enzyme in the absence of ligand, either the substrate or the metal. It is too early at this stage to definitively conclude the factors involved in contributing to the distinct specific activities, but may well be a direct consequence of the conformational differences between the well-separated samples. While most enzyme-catalysed reactions occur at a very fast rate, this result showed that it is not the case for the *Lm*iPGAM in P2 and P3. This may indicate major molecular and conformational changes of the enzyme structure during catalysis.

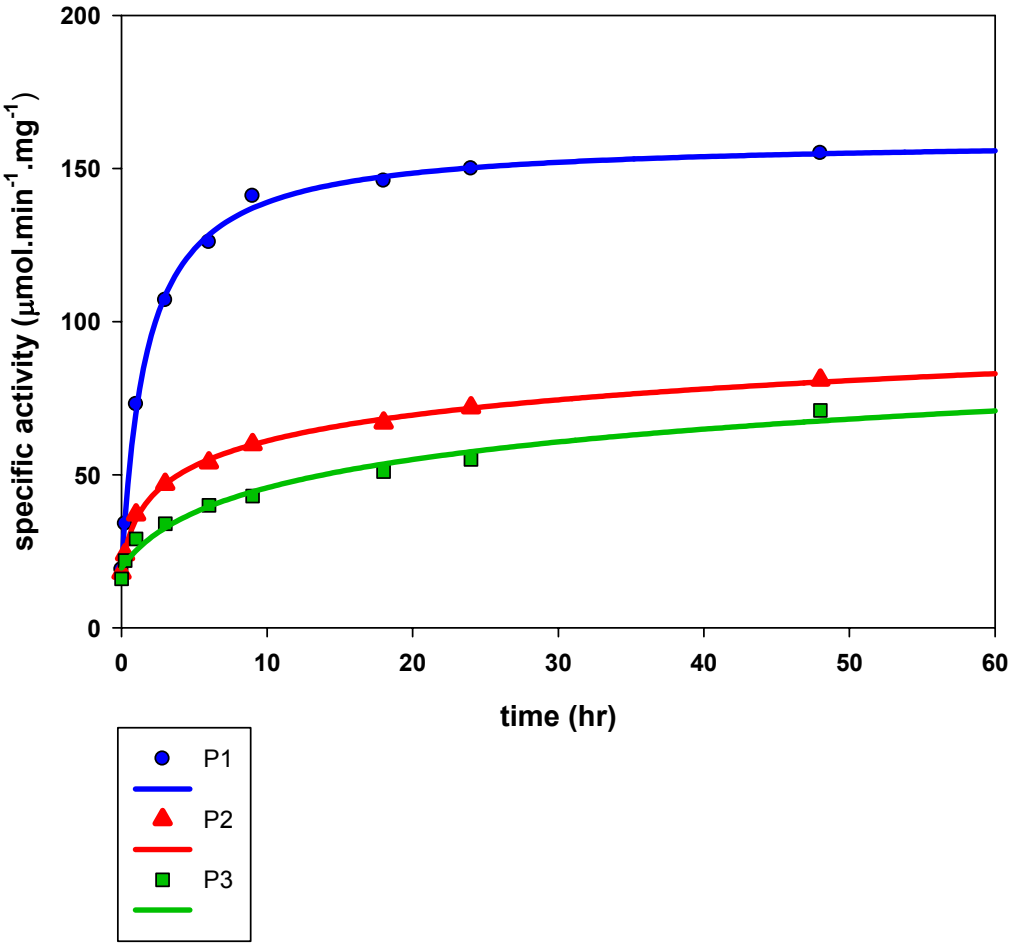


Figure 4.6 The specific activities of *Lmi*PGAM P1, P2 and P3 after incubation with 0.1 mM Co^{2+}. The specific activity was found to be similar prior to incubation, but was significantly enhanced during the incubation period, especially for P1. The maximum value for P1 is calculated to be 159.8 µmol.min^{-1}.mg^{-1}, whereas P2 and P3 maximum values are found to be off-scale. The off-scale refers to the slow pace of activity, which was too long and could not be measured in the period of experiment.

ii) Gel filtration

Gel filtration is based on the separation of proteins according to their sizes and shapes. This technique is not only utilised for the purpose of purification, but also as an analytical method for biophysical characterisation. In this study, a Superdex 200 10/30 GL chromatography column was utilised for the experiment, and the protocol was as explained in detail in chapter 2. This technique was anticipated to provide indications of the properties of the three peaks based on their estimated molecular masses. Using a calibration standard for molecular mass determination, in which the gel phase distribution coefficients (K_{av}) of the standard proteins were plotted against the logarithm of their respective molecular masses, (see Appendix I (c)), the molecular masses of the three peaks were estimated. The results indicated variations in the oligomeric states of *Lm*iPGAM, and also provided some indications of the existence of different conformations.

As depicted in Figure 4.7, P1 exhibited a single sharp UV elution peak, with an estimated molecular mass of 57.7 kDa and an elution volume of 14.14 ml, indicating the presence of a single monomeric species in the sample. P2 eluted at a slightly earlier volume compared to P1, resulting in elution at 13.98 ml. Interestingly, however, the molecular mass of P2, which is estimated to be 60.3 kDa also corresponds to a monomeric protein. These results indicate that both P1 and P2 may plausibly exist in monomeric states, but may exhibit different conformations, resulting in a delayed elution volume for P1. The apparent smaller conformation of P1 may indicate that the species exists in the closed form, assuming *Lm*iPGAM is in a normal, globular compact shape where the two domains are clasped close together resulting in a smaller radius, whereas P2 may well be in the open form. There is also a very small shoulder peak in P2 that eluted at 12.32 ml, in which the size may correspond to a dimeric *Lm*iPGAM. This observation thus suggests that P2 may well have a trace of dimer in the predominately monomeric sample.

The P3 sample exhibited two major peaks which were eluted at significantly different retention volumes. The first peak eluted at 12.16 ml, corresponding to a molecular mass of 127.8 kDa, whilst the second much larger peak eluted at around

the same retention volume as for P2: 13.95 ml, which corresponds to 63.2 kDa. These peaks again correspond to dimeric and monomeric states of *Lm*iPGAM, but with a higher proportion of the dimeric form compared to P2. Consequently, this observation further supports the idea that *Lm*iPGAM dimerises in solution, despite the monomeric nature of the protein obtained from the crystal structure.

It is therefore suggested that *Lm*iPGAM, by separation through the ion-exchange column after being initially purified by Ni^{2+}-IMAC, could be distinguished into different conformations and oligomeric states. These preliminary observations however, could be overestimated since the predicted molecular masses were relative to the calibration standards, and also might be concentration dependent. In the next few sections, further characterisation of the three peaks will be described, including size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS), which provides an absolute quantitation of the protein molecular masses.

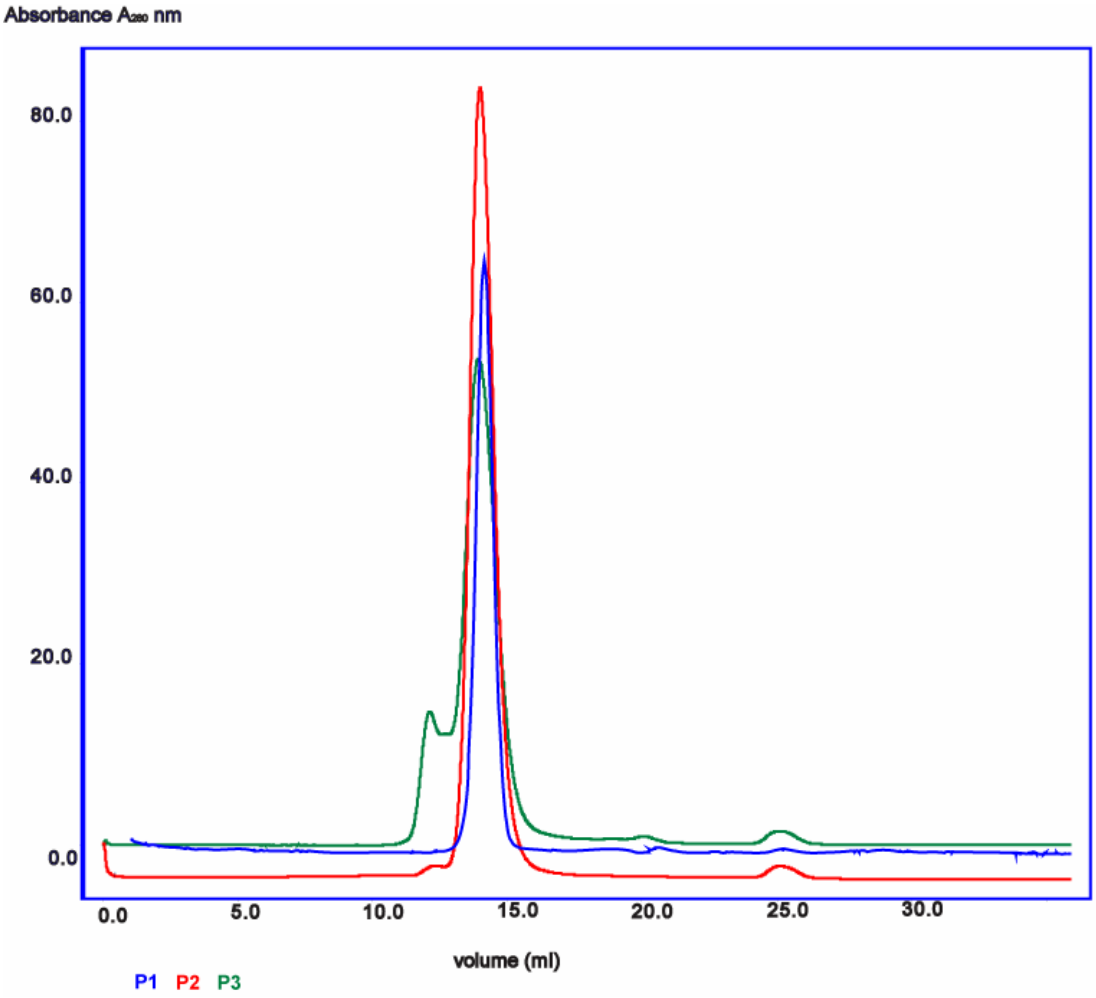


Figure 4.7 Gel filtration analysis of *Lm*iPGAM P1, P2 and P3. P1 was eluted as a single monomeric peak at 14.14 ml, which corresponds to the size of 57.7 kDa. P2 was eluted at a slightly earlier volume (13.98 ml) and corresponds to the size of 60.3 kDa. Both P1 and P2 are postulated to be monomeric proteins with different shapes and conformations. P3 on the other hand, was shown to be eluted as double peaks: the earlier peak was eluted at 12.16 ml, and estimated to possess a molecular weight of 127.8 kDa while the latter peak was eluted at 13.95 ml and corresponds to 63.2 kDa. P3 was thus predicted to contain a mixture of monomeric and dimeric proteins.

iii) Dynamic Light Scattering

Proteins in solution undergo constant random diffusion phenomena known as Brownian motion. This type of motion is advantageous for measuring the sizes and shapes of particles such as proteins in solution. Dynamic light scattering is one available method to measure the hydrodynamic size of proteins based on their diffusional properties. The size is frequently expressed in terms of the hydrodynamic diameter or the Stoke's radius (which corresponds to half the hydrodynamic diameter). It is important to note that these sizes correspond to the diameter of a hard sphere that diffuses at the same rate as the protein. These are inherently more extended molecules that are not perfectly spherical, and the hydrodynamic diameter of a protein is therefore less than its maximum rotational diameter.

In this particular experiment, the hydrodynamic diameters of the three *Lm*iPGAM peaks were determined in the absence of additives, and also in the presence of 1.5 mM 3PGA as the substrate and low (0.1 mM) and 1 mM (high) Co^{2+} concentrations. These conditions exhibit the highest activity, and are also the conditions used for the previous enzyme assay. These measurements were an important tool in predicting the characteristics of the three peaks, which interestingly differed from one another in terms of specific activity, as was discussed in section 4.5 i. The protein concentrations of all samples were standardised to 1 mg/ml, the measurements were repeated in triplicate, and incubation with additives were carried out for 1 hour at 25°C. The mode values of the hydrodynamic diameters of the protein molecules in P1, P2 and P3 are tabulated in Table 4.3a and b, and reported as both intensity and volume distributions (see below). In Figure 4.8 the hydrodynamic diameters (intensity distribution) for these samples in the absence of additives are shown. The results exhibit that the hydrodynamic diameters for all measurements by intensity distribution are larger than those of the volume distribution. This is because the intensity of the scattering is proportional to the diameter to the power of 6 (d^6), which is based on the Rayleigh approximation. Meanwhile, the hydrodynamic diameter by volume distribution is based on the volume of a sphere, which equals to $4/3\pi(d/2)^3$. It is noteworthy that the hydrodynamic diameter by intensity distribution is closer to the predicted values from the HYDROPRO software (see Table 4.3c), as

well as the measured maximal rotational diameter from the solved crystal structure (see section 4.4i). Hence, for the purpose of comparing the experimental and calculated values, only the hydrodynamic radius by intensity distribution was taken into account.

The results indicate that P1, which is postulated to be a monomeric closed form of *Lm*iPGAM is estimated to have a hydrodynamic diameter of 6.965 ± 0.184 nm by intensity distribution. The addition of the substrate 3PGA caused a small effect on the hydrodynamic size by intensity distribution, showing a decrease of only ~ 0.063 nm of hydrodynamic diameter, suggesting that this particular sample was not altered by the addition of 3PGA. The addition of both the substrate 3PGA and 0.1 mM Co^{2+} however, shifted the hydrodynamic size to a larger diameter (Table 4.3a). The addition of the substrate 3PGA to P2 resulted in a slight decrease of the hydrodynamic diameter by 0.117 nm compared to the native sample, although further addition of 0.1 mM Co^{2+} into the sample provoked a major increase of the hydrodynamic size. Peak P3, which had a larger proportion of material eluting at ~12 ml, showed a substantial shift of hydrodynamic diameter (0.786 nm), also indicating a smaller form of iPGAM with the addition of the substrate 3PGA. Nevertheless, in the presence of 0.1 mM Co^{2+}, the hydrodynamic diameter shifted to a larger value (see Table 4.3a).

These results suggest that the substrate 3PGA had a relatively small effect on the conformation or on the oligomeric state of the three samples. However, Co^{2+} at a concentration of 0.1 mM substantially increased the hydrodynamic size, although the mechanism is not totally understood. The hydrodynamic diameter differences between the closed-form *Lm*iPGAM and the open-form *Tb*iPGAM from the calculated values are ~0.40 nm (Table 4.3c), whereas the experimental closed and open-form of *Lm*iPGAM (P1 and P2) the diameter differences are found to be ~0.55 nm. Both sets of results indicate that the open-form is larger than the closed conformation, and agree with the previous crystallisation condition (Poonperm 2005) that reported the requirement of 3PGA in obtaining the closed form structure.

A second experiment was then conducted, where only P2 and P3 were analysed in the presence of 1 mM Co^{2+}. It was assumed that an insignificant effect would be observed on P1, as indicated in the first experiment, and was not included in this analysis. The results, as shown in Table 4.3b, showed that samples in the presence of 1 mM Co^{2+} exhibited a substantial change when incubated with 1 mM Co^{2+}. The reported hydrodynamic diameters were larger than the maximal rotational diameter of even a dimeric form of iPGAM, and perhaps correlated with the formation of higher order oligomers or of aggregated molecules.

Table 4.3 a) The estimated hydrodynamic diameters for *Lm*iPGAM P1, P2 and P3 in the presence and absence of substrate and low Co^{2+} concentration, incubated for one hour. The one hour incubation was performed based on the results from the previous specific activity measurement (see section 4.5i), which corresponds to the time to reach half of the maximum activity. Results are expressed in both intensity and volume distributions. b) Similar results are shown in the presence and absence of substrate and high Co^{2+} concentration. c) The comparison between the calculated values from HYDROPRO software and the experimental values.

a

Experiment 1 – low Co^{2+} (0.1 mM) **Samples**	**Mode value of hydrodynamic diameter (nm) by intensity distribution**	**Mode value of hydrodynamic diameter (nm) by volume distribution**
P1 native	6.965 ± 0.184	5.44 ± 0.147
P1 with 1.5 mM 3PGA	6.902 ± 0.95	5.613 ± 0
P1 with 1.5 mM 3PGA and 0.1 mM Co	7.239 ± 0.132	5.117 ± 0
P2 native	7.523 ± 0.202	5.882 ± 0.272
P2 with 1.5 mM 3PGA	7.406 ± 0	5.975 ± 0.314
P2 with 1.5 mM 3PGA and 0.1 mM Co	8.531 ± 0.787	6.452 ± 0.298
P3 native	8.787 ± 0.613	6.359 ± 0.444
P3 with 1.5 mM 3PGA	8.001 ± 0.212	6.368 ± 0.624
P3 with 1.5 mM 3PGA and 0.1 mM Co	9.77 ± 0.45	7.076 ± 0.327

b

Experiment 2 – high Co^{2+} (1 mM) **Samples**	**Mode value of hydrodynamic diameter (nm) by intensity distribution**	**Mode value of hydrodynamic diameter (nm) by volume distribution**
P2 native	8.251 ± 0.22	6.452 ± 0.29
P2 with 1 mM Co	29.17 ± 0.77	20.48 ± 0.945
P2 with 1.5 mM 3PGA and 1 mM Co	19.02 ± 2.08	10.85 ± 3.602
P3 native	9.569 ± 1.027	6.563 ± 0.49
P3 with 1 mM Co	29.68 ± 2.321	20.53 ± 1.896
P3 with 1.5 mM 3PGA and 1 mM Co	19.73 ± 1.328	13.67 ± 1.234

c

Conformations	**Experimental values (nm) (*Lm*iPGAM) by intensity distribution**	**Calculated values (nm) from HYDROPRO**
Closed-form *Lm*iPGAM	6.97 ± 0.184	6.42
Open-form *Tb*iPGAM	7.52 ± 0.202	6.82
Difference between open and closed conformations	0.55	0.40

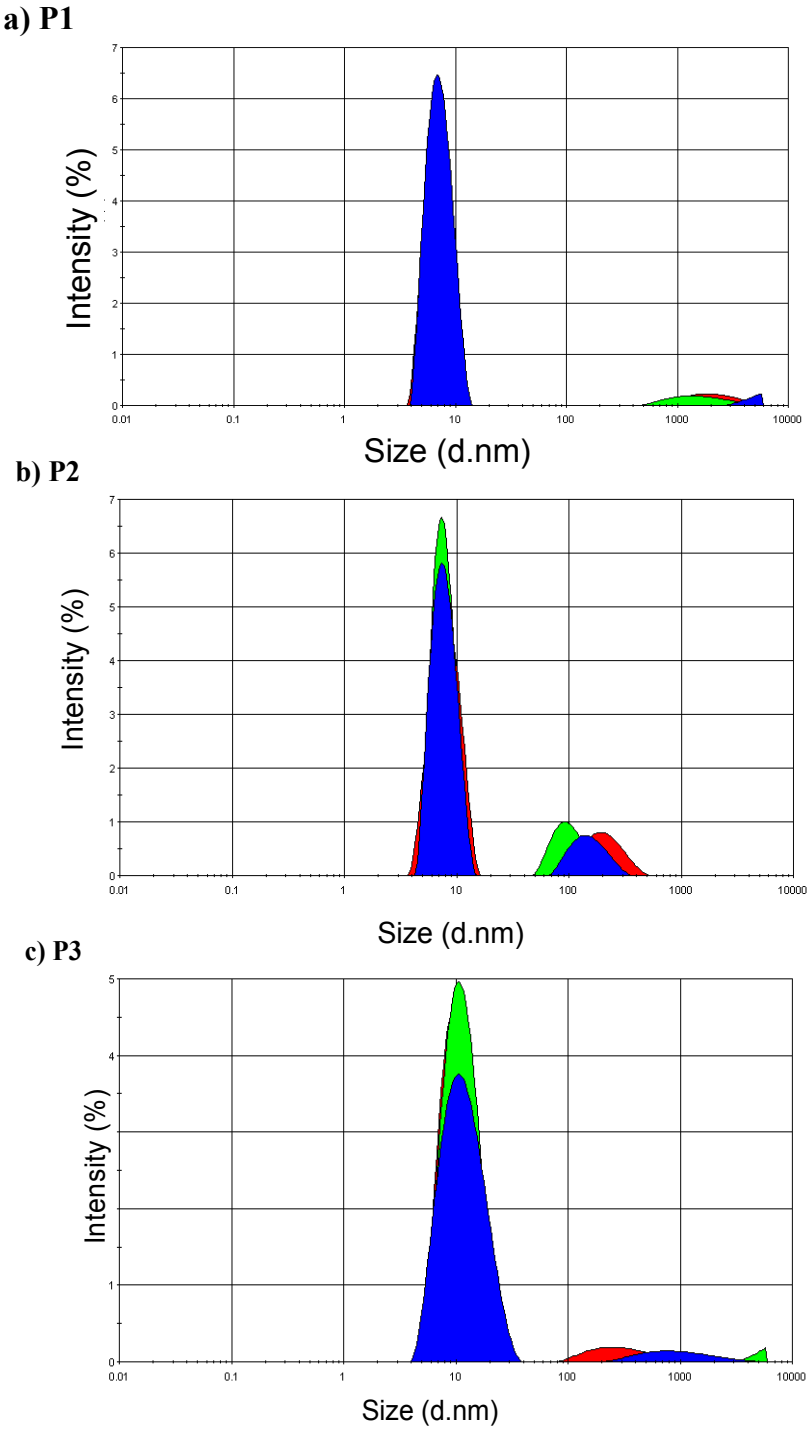


Figure 4.8: The hydrodynamic sizes of *Lm*iPGAM a) P1, b) P2 and c) P3 by intensity distribution, which is based on the relative intensity of the scattered light from the protein molecules in the samples been measured. All experiments were conducted in triplicate, and individual measurements are shown in **red** (replicate 1), **green** (replicate 2) and **blue** (replicate 3).

iv) Native-PAGE analysis

Polyacrylamide gel electrophoresis (PAGE) is a convenient technique for protein analysis, and is based on the uniform pore size of the gel. While SDS-PAGE is a technique used to separate proteins by molecular mass after treatment with SDS (see chapter 2), native-PAGE has the ability to separate proteins in the absence of SDS into their different conformations and surface net charges. However, it is often challenging to interpret the results unambiguously since the mobility of a protein may reflect both its conformation and surface net charge. Moreover, the success of the analysis also depends on the pI of the protein. Proteins with pI above the pH of the PAGE running buffer migrate backward or "retro-phorese" into the upper buffer chamber (http://www.nationaldiagnostics.com/article_info.php/articles_id/60). Hence, it is vital to ensure that the conditions for such analysis are appropriate for a particular protein.

The pI of His-tagged *Lm*iPGAM is 5.62, and the pH of the PAGE running buffer is 8.3, indicating that native PAGE could appropriately be employed to analyse the differences in the conformational or oligomeric states of the three samples: P1, P2 and P3. It is relevant to recall that all three samples were indistinguishable by SDS-PAGE (see chapter 3). Each peak, together with a sample of desalted protein without the ion-exchange step as a control (C) was loaded separately, with a protein amount of 5-10 μg per individual well. The protocol was explained in detail in chapter 2. The results, as indicated in Figure 4.9, showed that P1 is observed as a single prominent band, in contrast to P2 and P3 with additional multiple upper bands. P3 had a further additional band compared to P2. The upper bands in P2 and P3 may correspond to i) different conformations as a result of distinct charge distribution on the surface of the protein molecules, and/or ii) higher molecular masses due to different oligomeric states. It is indeed relevant to correlate the results with multiple conformations and/or oligomeric states of *Lm*iPGAM. The control (C) also showed the presence of multiple bands, in a pattern similar to that of P3, which indicates that the ion-exchange step was crucial for separating the *Lm*iPGAM into its different conformations and oligomeric states. In the next section,

SEC-MALS analysis will be discussed, as an additional powerful and complementary technique to characterise the samples.

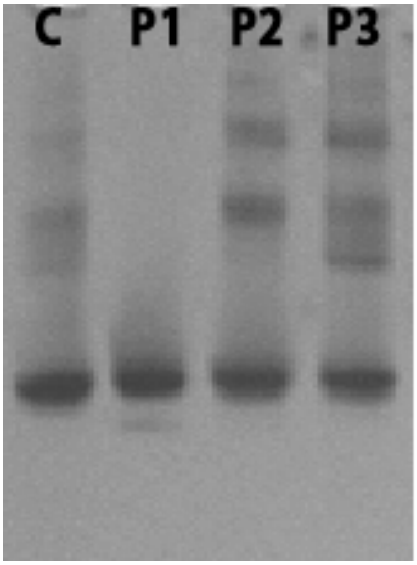


Figure 4.9 Native-PAGE analysis of *Lm*iPGAM P1, P2 and P3. A 9% native-PAGE is shown with 5 -10μg samples of each protein fraction. C is a control sample of pure *Lm*iPGAM without an additional ion-exchange step.

4.6 Size-Exclusion Chromatography coupled with Multi-Angle Light Scattering (SEC-MALS)

i) Different oligomeric states are confirmed by SEC-MALS

The downside of molecular mass determination from a conventional gel-filtration analysis is the reliance of the technique on calibration curves, in which relative comparisons are being made with a series of globular protein standards. By combining size-exclusion chromatography (SEC) to static multi-angle light scattering (MALS) or Rayleigh scattering, calibration standards are not required, and a direct measurement of the molecular mass can be independently employed. Following the earlier molecular mass determination by gel filtration, it was relevant to determine the absolute molecular mass of the three peaks: P1, P2 and P3 in order to confirm the presence of different conformations and/or oligomeric states of

*Lm*iPGAM in solution. The experiment was conducted for the three samples, with a standardised concentration of 0.4 mg/ml for each peak (Figure 4.10). P1 exhibited a single peak eluting at ~13.9 ml with a calculated molecular mass of 58.23 kDa, which corresponds to monomeric *Lm*iPGAM, and is smaller than the mass calculated from the sequence, 61.8 kDa. The peak was shown to be homogenous, and monodispersity was observed, indicating the presence of a single species in the solution. The peak was predicted to be in the closed conformation, and will be discussed in section 4.6iii.

P2 exhibited a major peak eluting at ~13.4 ml, corresponding to a monomer at 59.1 kDa. In addition, the small peak eluting at ~11.6 ml had a molecular mass of 114.9 kDa, indicating the presence of a dimeric form of the enzyme. The heterogeneous protein eluting before 11 ml may correspond to a mixture of higher order oligomers and aggregated proteins. The major peak was slightly asymmetric with a right-hand edge, indicating that the *Lm*iPGAM in this peak might contain a small amount of the closed form in addition to the predominantly open form. P3 clearly contained multiple forms of iPGAM, with a peak at ~11.6 ml, another peak at 13.4 ml and a shoulder at 13.9 ml, in addition to heterogeneous protein eluting at ~9.7 ml. The *Lm*iPGAM molecules eluting at 13.4 ml and 13.9 ml were monomeric (56.88 kDa), presumably in the open and closed forms, respectively. The *Lm*iPGAM eluting at 11.6 ml had a mass of 113.7 kDa, corresponding to a dimer, and the protein eluting at ~9.7 ml correspond to a tetramer (265.5 kDa). However, the vertical line corresponds to the molecular mass was broad, which may indicate that apart from a higher order oligomer, it may also contain aggregates (see Figure 4.10).

Strong justification for the existence of multiple conformations and oligomeric states of *Lm*iPGAM has been obtained, and this interestingly, correlates well with the earlier gel filtration and native-PAGE experiments. The results from SEC-MALS analysis are useful to determine more accurately the absolute molecular masses, and the results tally with the gel filtration results. Although the concentrations of the samples were dissimilar in these two analyses, they exhibit comparable results, in which the molecular mass in the gel filtration analysis was determined relative to protein calibration standards. This observation strongly

suggests that trypanosomatid iPGAM, although previously reported as a monomer, can also be present in a dimeric state in solution, as indicated by the properties of P2 and P3. These data also equated with the native PAGE results, which display the presence of upper bands in both samples, with P3 having more bands compared to P2. The additional upper bands may well be the consequences of sample aggregation, as also observed from the SEC-MALS analysis.

The phenomenon of dimer formation from monomeric proteins has previously been discussed (Park and Raines 2000; Marianayagam et al. 2004; Ali and Imperiali 2005). The occurrence of such observations can be related to a domain swapping mechanism, where a monomeric protein reciprocates one of its domains with another segment of a monomer to form a homodimer (Figure 4.11), before residues been replaced to stabilise the dimeric form (Park and Raines 2000). *Lm*iPGAM in this case however, exhibits the ability to exist both as a monomer and dimer in solution. This appearance has also been reported for some proteins that are known to exist in both the monomeric and oligomeric states, in which the interdomain interactions can occur either within a sole monomer or between monomer units (Liu and Eisenberg 2002; Ali and Imperiali 2005). A significant presence of polar and charged residues at protein intersubunit interfaces, which contributes to distinct surface charges in the two domains, may also encourage the formation of a dimer in monomeric proteins (D'Alessio et al. 1999). In this case, STP, a program which possesses the ability to predict protein-protein or protein-ligand interactions by the molecule surface properties (Mehio et al. 2010) has been utilised for the purpose of predicting the most likely region to be involved in the interactions. The results, as shown in Figure 4.12, indicate the regions which are most likely to interact with other biomolecules (red), whereas the blue- coloured regions are less likely to form interactions. Nevertheless, the areas which are prone to form protein-protein interactions with another *Lm*iPGAM molecule may well be spread out through the whole top and bottom regions of both the phosphatase and transferase domains. This may result in instability of the dimer molecule if it involves weaker interactions, as coloured in green or blue. Intriguingly, from the SEC-MALS data, there seems to be a dynamic equilibrium between a monomer and

dimer in *Lm*iPGAM, where the presence of both states could be observed in P2 and P3 (Figure 4.10). It is indeed relevant to investigate this phenomenon in the presence of divalent metal ions in each sample, as this may contribute to the shift of equilibrium towards either a monomer or dimer. It is noteworthy also, that it is unlikely for the dimer formation to be mediated by disulfide bridges, because the five thiol (SH) groups are buried in the protein and not available to make such interactions.

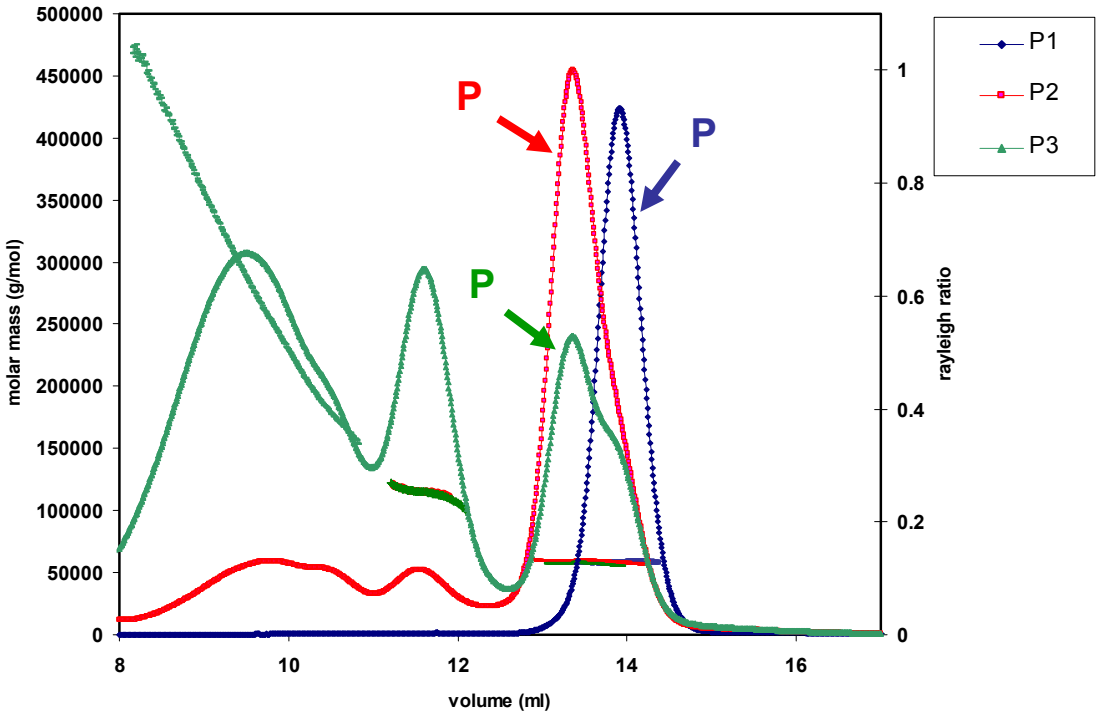


Figure 4.10 SEC-MALS analysis of *Lm*iPGAM P1, P2 and P3, where the presence of monomeric species was indicated in the major peaks eluting between 13-15 ml. P2 and P3 also exhibited the presence of dimeric species in the peaks eluting at ~11.6 ml. A tetramer was observed in P3 in the peak eluting at ~9.7 ml. The molecular masses are shown as the horizontal lines corresponding to the colours of each peak. The essentially horizontal lines for the peaks at 13-15 ml indicate that the masses are consistent throughout the elution peak. However the lines for the peaks at 11.6 ml and 9.7 ml have vertical components indicating changes in molecular mass across the elution peak. The peak eluting at ~9.7 ml may contain a mixture of tetramer and aggregated materials.

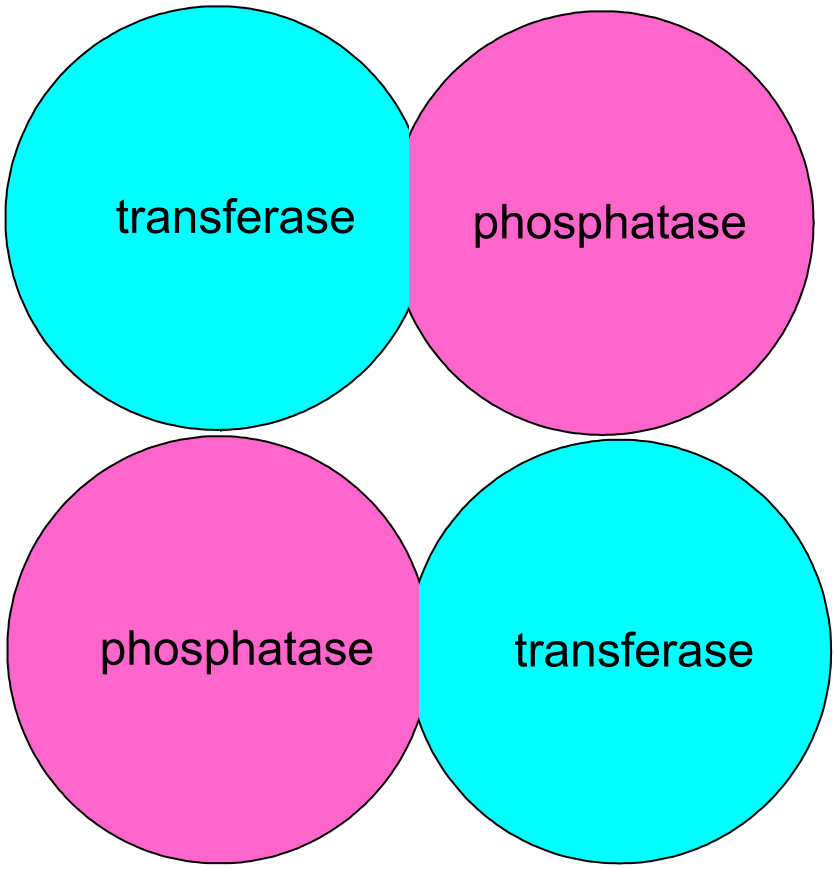


Figure 4.11 The formation of dimeric *Lm*iPGAM by a domain swapping mechanism.

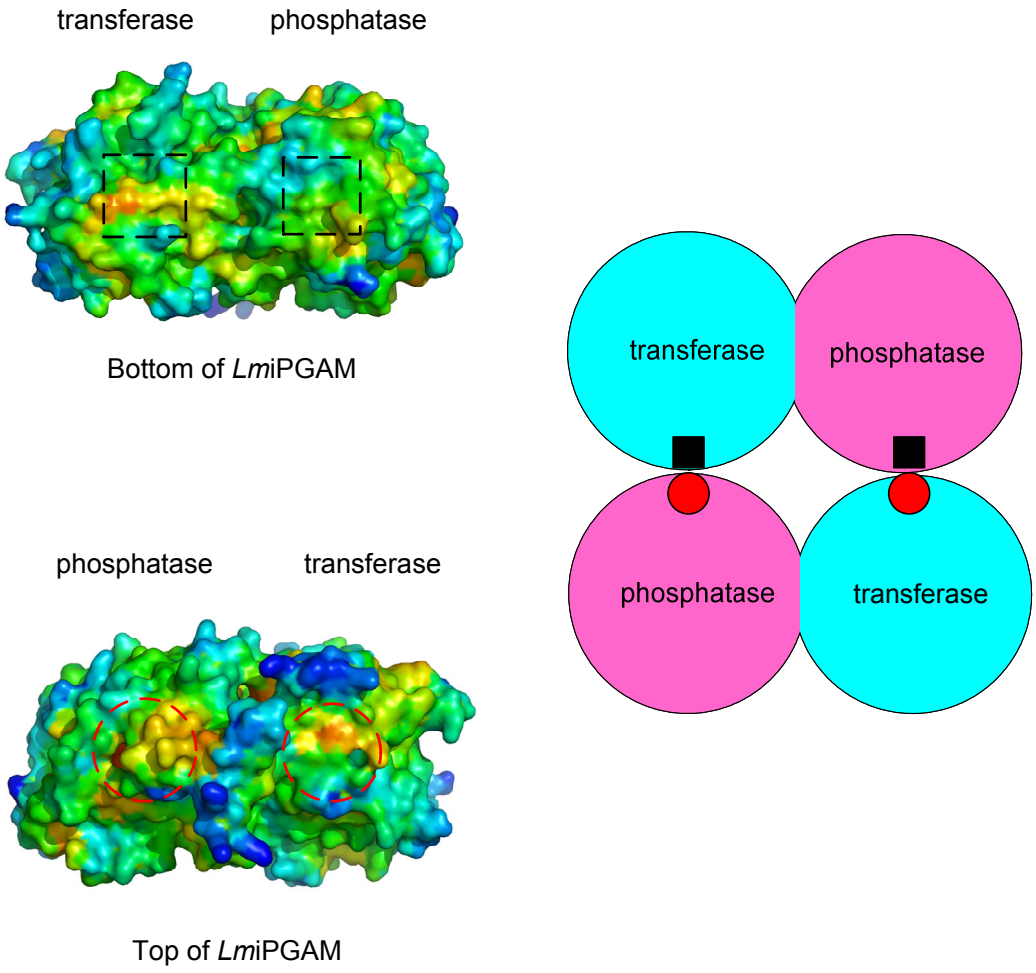


Figure 4.12 STP analysis on the closed-form *Lmi*PGAM (PDB ID 3IGY). The red regions represent the most likely areas for interactions with another molecule, whereas the blue regions are the least likely to form any inter-molecular interactions. The red circle represents the top of the phosphatase and transferase domains from one *Lmi*PGAM molecule, which is likely to interact with the bottom of another *Lmi*PGAM molecule (indicated with the black box). The formation of dimer may involve weaker interactions in the protein interface, which are coloured in green or yellow.

(ii) The effects of Co^{2+} on *Lm*iPGAM oligomeric states

The presence of a dimeric form of iPGAM, in addition to the monomeric form has been determined in both P2 and P3 samples, as discussed in detail in the previous section. P1 however, consistently displayed the presence of only a single monomeric state in the sample. In the previous DLS experiment, the effect of 1 mM Co^{2+} was analysed in both P2 and P3, where the hydrodynamic diameter shifted into much larger sizes, which may indicate the presence of aggregated material in the sample. Although a similar treatment was not tested on P1, it would be interesting to observe if the presence of metal could prove likewise for P1. Hence, in the next SEC-MALS experiment, the three samples were incubated with 1 mM Co^{2+} for 1 h, before being analysed sequentially for reproducibility.

Figure 4.13 shows that P1 eluted predominantly as a monomeric peak at 14.3 ml, with perhaps surprisingly a trace of dimer (not present in the absence of added Co^{2+}) at 12.3 ml. The molecular mass of the monomeric form of the enzyme corresponds to 65.8 kDa, and the molecular mass of the dimeric form was found to be 126.4 kDa. The molecular mass of the monomeric P1 in the absence of Co^{2+} (see earlier section) was found to be 58.6 kDa, which was 7.2 kDa smaller than the value after incubation with 1 mM Co^{2+}. P2 and P3, by contrast, each exhibited single peaks, which corresponded to monomeric species of 64.6 kDa and 71.4 kDa, respectively, although the quality of the P3 data had become poorer.

It can be seen in summary that the incubation with 1 mM Co^{2+} has caused all three samples to adopt monomeric forms. It is also a striking observation that in the presence of Co^{2+}, the molecular masses of all the monomeric peaks have increased, and the elution volumes were shifted to be slightly delayed, thus suggesting that it is possible that Co^{2+} caused iPGAM to become more compact. It is noteworthy however, that the sizes of the peaks which were eluted from the column in the monomer position (see Figure 4.13) were diminished for P2 and P3, compared with the result in the absence of Co^{2+} in Figure 4.10. This might be due to a large amount of aggregated material, which was eluted near the void volume (8 ml). Nevertheless, all samples exhibited similar molecular masses, although it is apparent that the

molecular mass of P3 was slightly inconsistent throughout the peak, indicating the heterogeneity of the sample. The formation of a small amount of dimer in P1 however, is yet to be further analysed in the future. Nevertheless, the phenomenon of dimer formation caused by divalent metal ions has been discussed in other enzymes, such as *β*-Lactamases from *Pseudomonas aeruginosa*. It was reported that the dimeric form of this enzyme has been stabilised by a number of divalent metal ions (Danel et al. 2001). It is noteworthy that a similar experiment was attempted in the presence of Mg^{2+} and Zn^{2+} separately for *Lm*iPGAM P2, but the protein precipitated upon the addition of these metals.

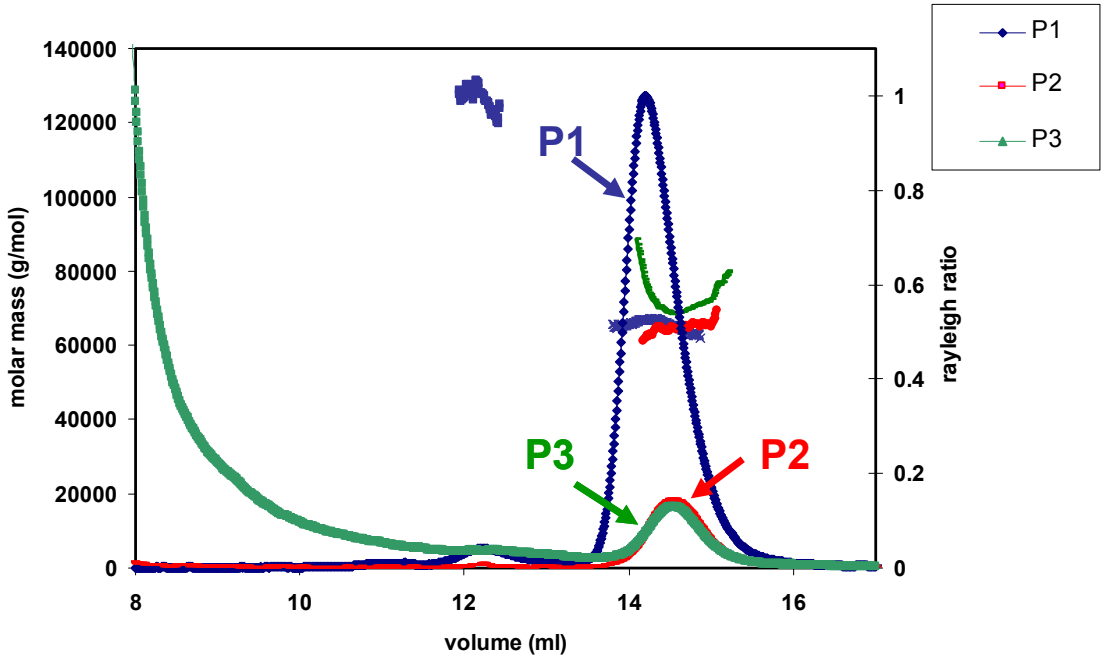


Figure 4.13 SEC-MALS analysis of *Lm*iPGAM P1, P2 and P3 in the presence of 1 mM Co^{2+}, where the presence of monomeric species was indicated in the distinct major peak. P1 surprisingly exhibited the presence of dimeric species, as indicated in the graph.

(iii) SEC-MALS analysis provides strong indications of the different conformational states of *Lm*iPGAM in solution

With regard to understanding the potential conformational differences displayed by *Lm*iPGAM, significant clarification has been obtained by comparing the SEC-MALS elution volumes of the three samples, thus giving an indication of contrasting macromolecular sizes or shapes. This had been observed earlier in the gel filtration experiment, where the elution volume of P1 was delayed compared to the other two peaks, presumably corresponding to a smaller hydrodynamic size of the protein molecule. Based on this hypothesis, SEC-MALS analysis was employed and repeated in triplicate, to provide reproducible observations regarding the elution volume of the peaks.

In the first run of the experiment (black lines in Figure 4.14), which was carried out 24 h after protein purification, the elution volume of P1 was 13.95 ml (Figure 4.14a). A symmetrical peak was observed, with no apparent leading or trailing edges. The major peak of P2 that corresponds to a monomer, on the other hand, was eluted at 13.38 ml (Figure 4.14 b), indicating that the molecule has a larger hydrodynamic size, and hence resulting in an earlier elution volume. This sample however, showed a distinct right-hand edge towards the elution volume of P1. The third sample (P3) exhibited a similar elution volume to P2 (13.39 ml) for the monomeric protein, but interestingly, the presence of the right-hand edge of the shoulder peak was more apparent (Figure 4.14 c).

The second run of the experiment was conducted 43 h after purification, where similar patterns of elution were observed. A single symmetrical peak was observed for P1 with no apparent leading or trailing edges. However, the shoulder peak of the monomeric P2 became more discernible, and an apparent shift was observed towards the elution volume of P1. A similar observation was seen for the monomeric P3, where the distinct shoulder peak became much clearer compared to the sample that had been analysed in the first batch.

After 50 h, the last run of the experiment was conducted, and as observed in Figure 4.14 a, a single peak was observed for P1, this time with a slight left-hand

edge. The shoulder peak of monomeric P2 was observed to be very similar to the shape of the monomeric P3 from the previous two batches. By this time, the monomeric P3 exhibited two distinguishable monomeric peaks, where the elution volume of the first distinct peak corresponds to the elution volume of P2 and P3 (13.42 ml), and the second peak corresponds to that of P1 (13.92 ml). One of the most important observations from this analysis is the consistency of the elution volume of P1 and the monomeric peaks in P2 and P3. The difference in the elution volumes was ~0.5 ml for the two monomeric peaks (P1 and P2/P3, ignoring the existence of shoulder peaks). This was also observed in the earlier gel filtration experiment, which showed that the open and closed conformations can be well-separated by ~0.5 ml difference in the elution volume.

It was postulated at the beginning of this chapter, that trypanosomatid iPGAM can either be in the closed, or the open conformation. A direct understanding from the SEC-MALS data is that *Lm*iPGAM, as a globular protein, exists in both conformational states, and this can be seen from the difference in the elution volumes of the monomeric peaks of all samples. P1 may well be entirely in the closed monomeric form, whereas P2 and P3 can be a mixture of populations: a dimer as well as a monomer with an open conformation. Interestingly, the monomeric peaks of P2 and P3 both have right-hand edges towards the elution volume of P1, which became more apparent when left longer in solution. This observation may indicate the presence of both the open and closed conformations in P2 and P3, and thus suggests that the change of conformation is time-dependent and could attain equilibrium after having been left for a certain period of time. This may well be the reason for the slow increase in reaching maximal activity of P2 and P3, compared to P1 that has been discussed in an earlier section (see section 4.5i). Nevertheless, both samples also exhibit a tendency to form aggregates, especially P3, where the intensity of the scattered light detected larger macromolecules in the earlier elution volume, corresponding to the void volume (8 ml). Interestingly, the formation of tetramer was detected in the first batch of P3 at around ~9.7 ml, which suggests that the peak probably contained not just denatured protein, but also higher order oligomers. These may also explain the difficulties in obtaining crystals for both P2

and P3, because of the presence of multiple species, as well as the formation of aggregated materials.

Open and closed conformations of the same protein have been widely reported (Ruzheinikov et al. 2004; Nallamsetty and Waugh 2007; Kumar et al. 2010). It is predicted that the open conformation of *Lm*iPGAM may be very similar to the open conformation of *Tb*iPGAM, as discussed previously (see section 4.3ii), since both enzymes share ~78% sequence identity. It is also relevant to postulate that the enzyme can possibly exist in a mixture of both the open and closed conformations (Bucher et al. 2011). In this case, the presence of ligands which induced the closed conformation would play a more important role in shifting the equilibrium towards the closed conformation. This mechanism is known as the 'conformational selection' or 'population shift' phenomenon (Boehr et al. 2009; Bucher et al. 2011). It has been proposed as well (Bucher et al. 2011) that the transformation between the forms may well be the consequence of solvation and would be assisted by the packing of non-polar side chains. If the conformation is different, and the surface charge of the macromolecule consists of more non-polar residues, it may well affect the hinge region, which is important in allowing the opening and closing motion to occur (Hayward 1999). It is thus suggested that the presence of alternative conformations in *Lm*iPGAM may well be related to the difference in surface charge of the molecules, and can change occasionally from one conformation to the other, as observed in P2 and P3. Structural changes of *Lm*iPGAM are postulated to be important in balancing the conformational dynamic equilibrium for protein integrity and functionality.

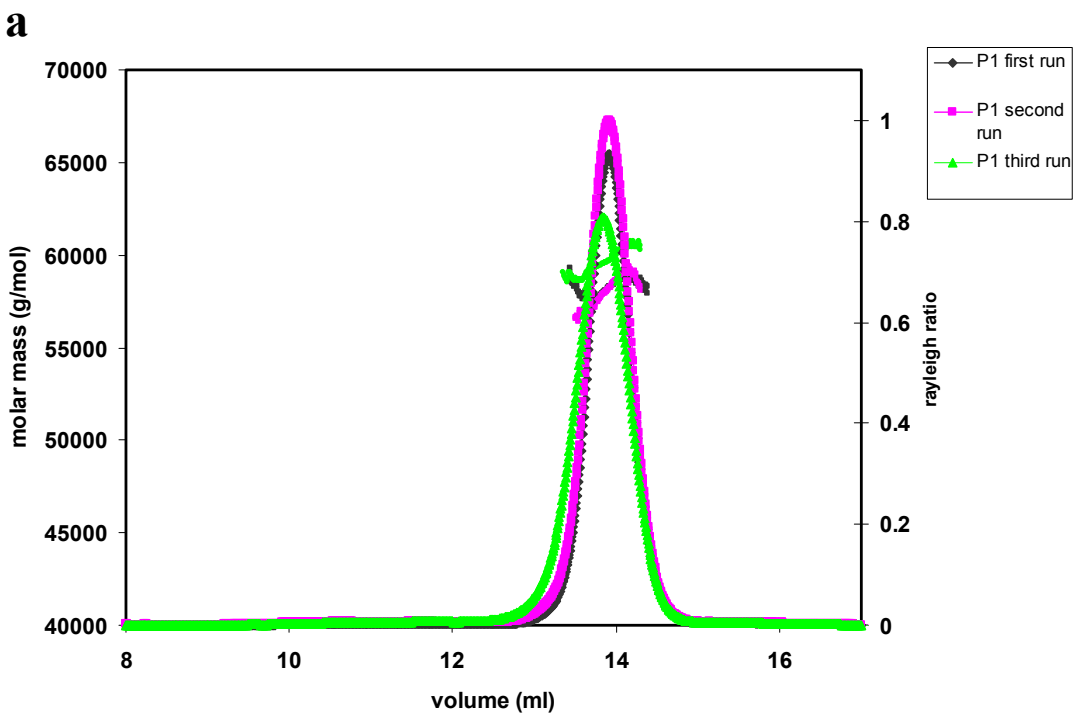
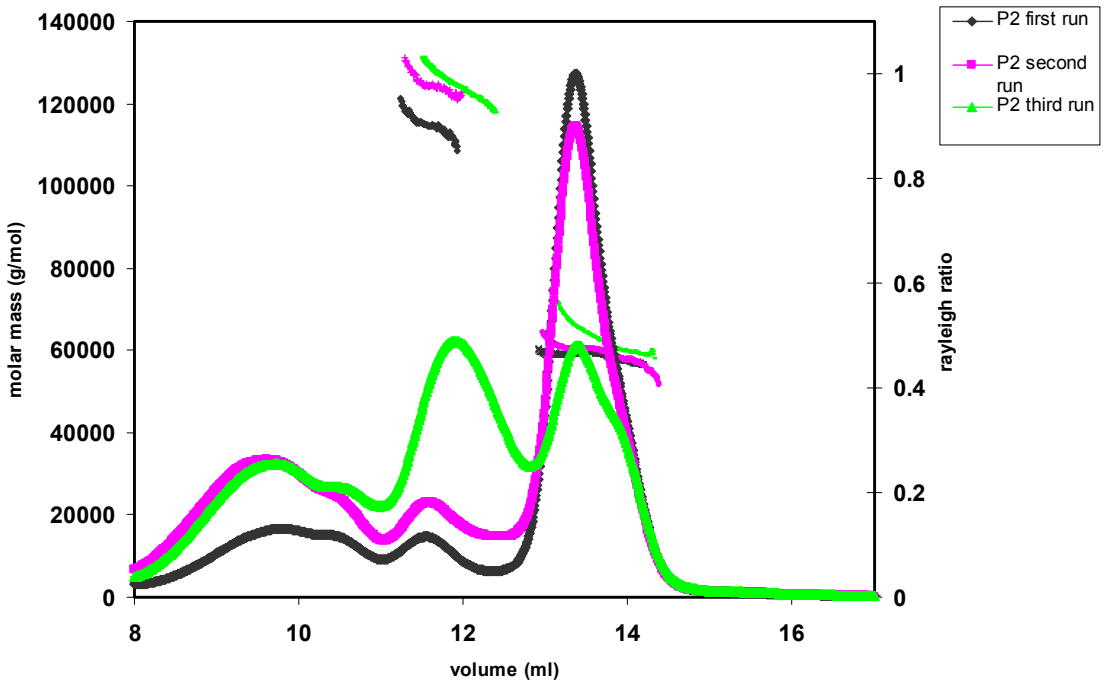


Figure 4.14 SEC-MALS analysis in all *Lmi*PGAM samples as observed in three separate runs: a) P1, b) P2 and c) P3. The mixture of aggregated materials and higher order oligomers were observed in all samples, but became more apparent in run 2 and run 3.

b



c

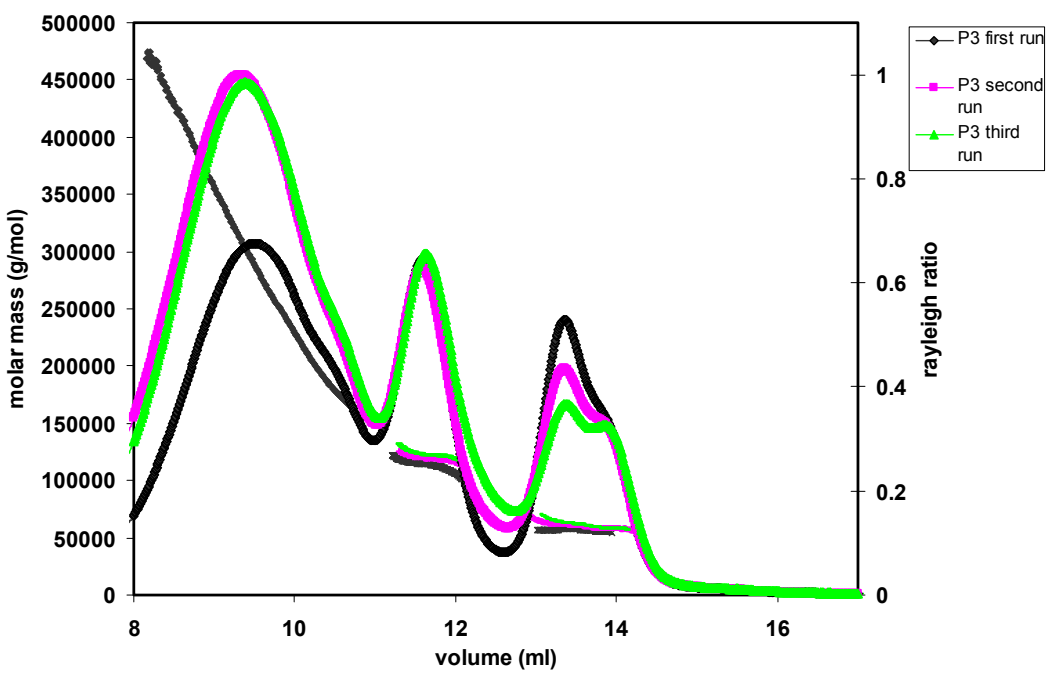


Figure 4.14 contd.

4.7 Different forms and different oligomeric states can be separated by ion-exchange chromatography

A significant outcome from the SEC-MALS analysis, as well as the earlier experiments, is the demonstration that an ion-exchange step is capable of separating and consequently capturing different conformational and oligomeric states of *Lm*iPGAM. This can be explained by the anionic nature of the column used, MonoQ 5/50 GL (GE Healthcare). The positively charged column will strongly interact with negatively charged groups distributed on the surface of the *Lm*iPGAM. Different conformations or oligomeric states of the same protein possess different surface charges which affect the elution of each state as it goes through the column. Molecules with more positive surface charges will have weaker interaction with the column, and hence will be eluted first. The more negatively charged the molecule, the stronger the interaction with the column. A schematic diagram showing the interaction between the surface charges of the column and *Lm*iPGAM is illustrated in Figure 4.15.

In the case of *Lm*iPGAM, it is apparent that P1 has more positive charges on the surface of the protein, and this resulted in an earlier elution volume, compared to the other two peaks (see chapter 3). Subsequently, P2 was eluted, followed by P3, which indicates that the binding between the protein molecules and the column is tighter than P1. Factors that contribute to this observation may well be a) the differences in conformation and oligomeric states, which result in different surface charges, and/or b) the presence of additional positively charged molecules, such as metal ions, in the early eluting fractions. Nevertheless, the different elution volumes of the three distinct peaks P1, P2 and P3 have clearly correlated with significant differences in the conformational and oligomerisation states of *Lm*iPGAM, and ion-exchange has been found to be a useful method in separating the distinct protein molecules.

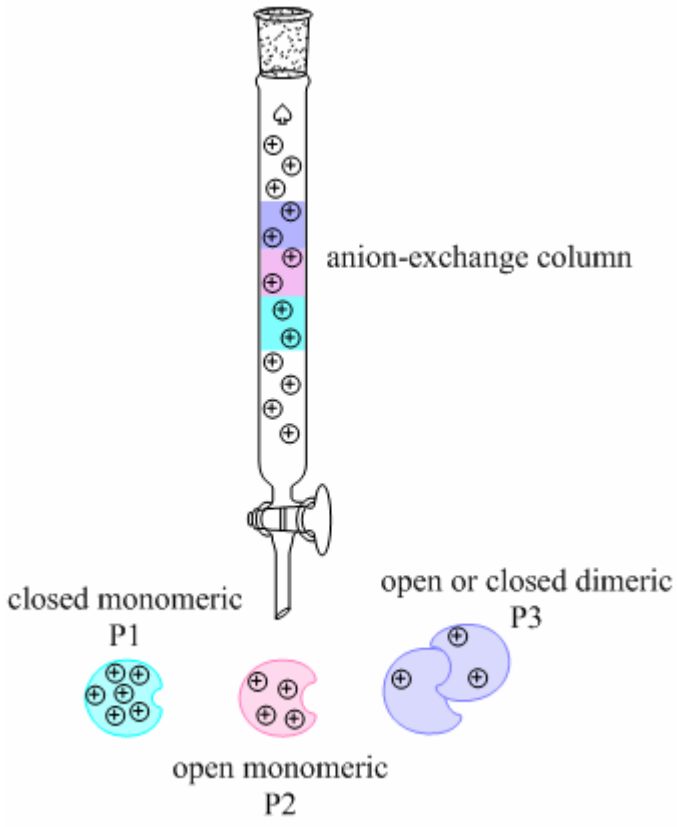


Figure 4.15 The interactions of various forms of *Lm*iPGAM with the anion-exchange column resulting in different elution volumes for each form.

4.8 Metal binding and metal content in *Lm*iPGAM

The previous section has discussed the utilisation of an anion-exchange column, MonoQ 5/50 GL in separating *Lm*iPGAM based on its different conformations and oligomeric states. It is also relevant to note that the presence of different positively charged ligands such as metal ions could also be in accordance with the different elution volumes from the anion-exchange column. Hence, the three samples P1, P2 and P3 were analysed using matrix-assisted laser desorption ionisation time-of-flight (MALDI-TOF) mass spectrometry, as well as ICP-MS analysis.

i) MALDI-TOF mass spectrometry of ion-exchange peaks from *Lm*iPGAM

Differences between P1, P2 and P3 of *Lm*iPGAM were further characterised by MALDI-TOF mass spectrometry of each individual sample. Samples of 0.5 µl of 1 mg/ml protein were used in this experiment, in the presence of 20 mM TEA pH 7.6 and 50 mM NaCl. Since this analysis was employed to analyse the whole protein molecule, and iPGAM is catagorised as a large protein molecule (~60.7 kDa), it would be challenging to differentiate the contributions of bound ligands to the exact mass of each peak. Nevertheless, the MALDI-TOF analysis could provide an indication if there are more metal ions or other ligands bound to one sample than another.

The results in Figure 4.16 a, b and c show MALDI-TOF spectra with a distinct major peak in all samples, indicating the presence of a monomeric molecular ion corresponding to *Lm*iPGAM. The masses of the peaks were found to be very similar to the calculated theoretical value of 61788.5 Da, including the His-tag and the initiator methionine. P1 was shown to be 61817.16 Da, whereas P2 and P3 were estimated to possess masses of 61762.81 Da and 61802.58 Da, respectively (Table 4.4). The augmented masses of P1 and P3 are postulated to be the consequences of Na^+ (molecular mass of 22) or Mg^{2+} (molecular mass of 23) bound to iPGAM (a Na^+ and Mg^{2+} ion in place of H+ in positive ion mode resulted in the increase in molecular masses). The small differences of the molecular mass however, may be overestimated, since the size of iPGAM is too large to be accurately determined by MALDI-TOF analysis. It is also noteworthy that apart from the main peaks that correspond to a monomeric protein, smaller peaks with masses of 30864.36, 30863.23, and 30858.90 Da were also observed in samples P1, P2 and P3, respectively. The identity of the smaller peak may well be of the doubly-charged molecular ions.

Table 4.4 The theoretical and experimental values of *Lm*iPGAM P1, P2 and P3 molecular masses

Samples	Theoretical (Da)	Experimental (Da)	Difference (Da)
P1	61788.5	61817.16	28.66
P2	61788.5	61762.81	-25.69
P3	61788.5	61802.58	14.08

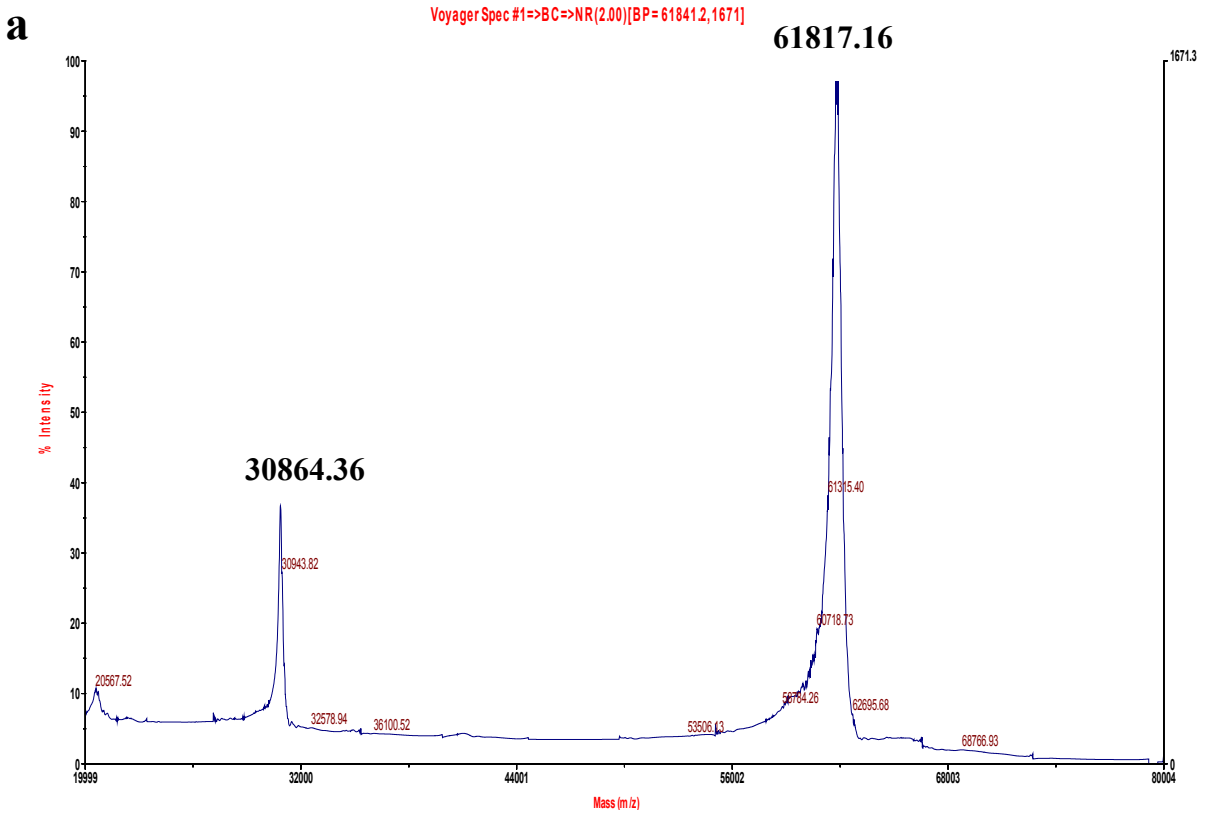


Figure 4.16 MALDI-TOF mass spectrometry of *Lm*iPGAM (a) P1, (b) P2 and (c) P3. P1 was shown to be 61817.16 Da, and P2 and P3 were estimated to possess masses of 61762.81 Da and 61802.58 Da, respectively. The augmented masses of P1 and P3 are postulated to be the consequences of Na^{+} (molecular mass of 22) or Mg^{2+} (molecular mass of 23) bound to iPGAM, resulting in small increases of the molecular mass of each sample.

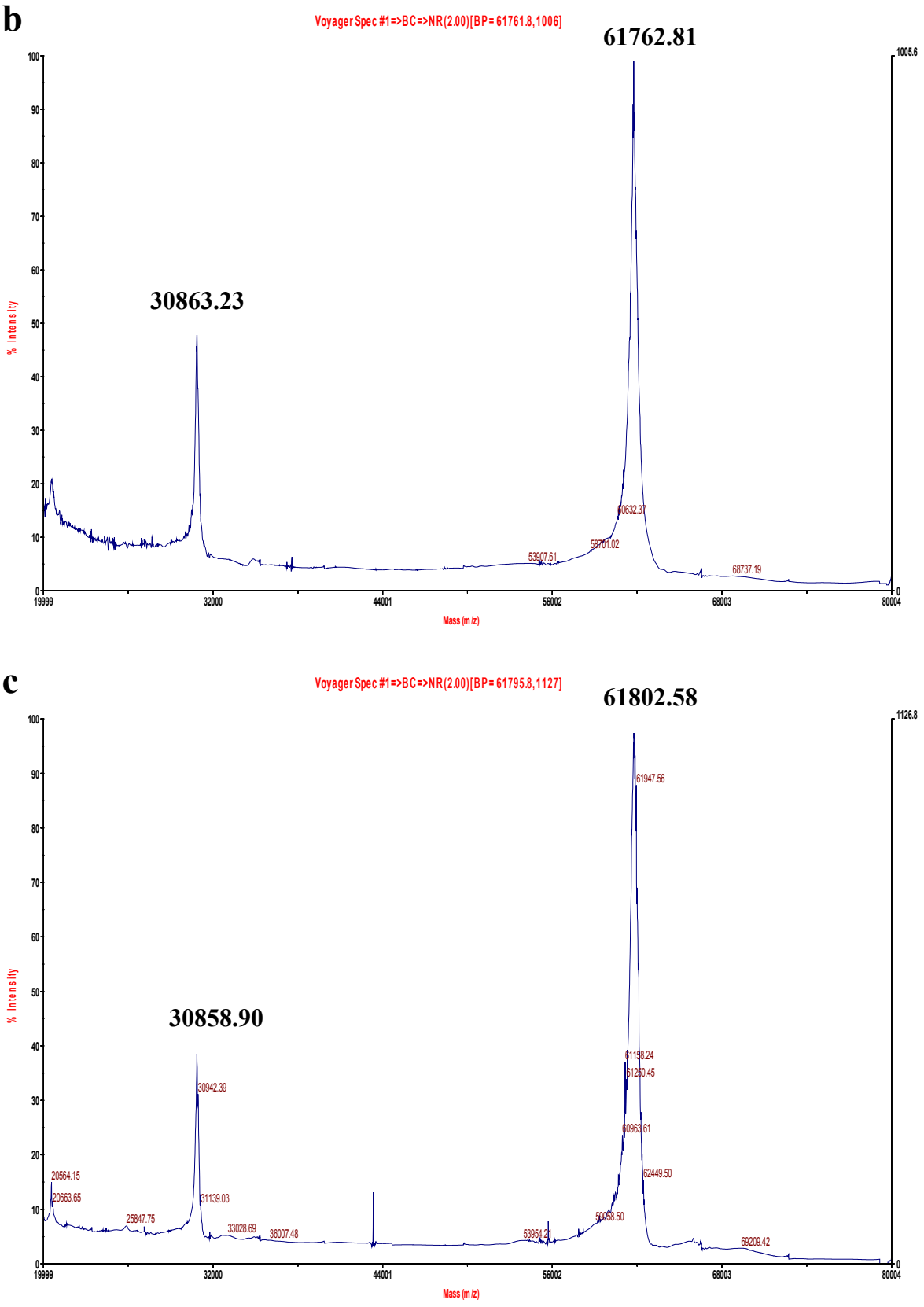


Figure 4.16 contd.

ii) ICP-MS of ion-exchange peaks from *Lm*iPGAM

The effects of divalent metal ions to *Lm*iPGAM has been discussed in earlier publications (Guerra et al. 2004; Poonperm 2005; Nowicki et al. 2009), where metal ions were reported to be essential for enzyme activity. However, the issue of which divalent metal ions that can support trypanosomatid iPGAM activity has always been controversial, for which the determination of metal concentrations in the protein samples are vital. In this case, ICP-MS analysis is one of the most appropriate methods for metal determination in purified protein samples, where the results will provide the concentrations of the divalent metals, thus providing clues of which metals are present at sufficient concentration to support enzyme activity, either *in vivo* or *in vitro*. This technique however, can not distinguish between free or bound metals. This interesting topic will be discussed in chapter 5.

Following on from the MALDI-TOF mass spectrometry analysis, which showed an apparent increase in the molecular mass of P1 and P3 of *Lm*iPGAM, ICP-MS was employed to measure the concentrations of divalent metal ions in the three samples P1, P2 and P3. This result, apart from being able to provide knowledge about the individual metal concentrations in each sample, would also indicate samples that had acquired more metal ions than the others. In this analysis, each sample of *Lm*iPGAM was normalised to 1 mg/ml, and seven metals (Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}) were measured, following the experimental procedure that has been described in chapter 2. As mentioned in chapter 2 and chapter 5, elemental standards were prepared and provided linear concentration-response curves in both water and in enzyme storage buffer (20 mM TEA-HCl, pH 7.6 and 50 mM NaCl) (see Appendix I (e)). The LOD for each individual metal has also been determined, in both water and the enzyme storage buffer.

The results are shown in Figure 4.17 a and b, where the metal concentrations in the three samples were divided into two categories: the more abundant and less abundant metals. Mg^{2+} was shown to be the most abundant metal in P1 (7.55 μM), whereas the concentrations of Mg^{2+} in P2 (0.55 μM) and P3 (0.52 μM) are ~14-fold lower. Zn^{2+} was also present at significant concentrations, with a concentration of

5.55, 0.8 and 3.26 μM in P1, P2 and P3, respectively. Fe^{2+} exhibited lower concentrations in all three samples compared to the other two abundant metals, nevertheless displaying a higher concentration in P1, rather than the other two peaks. The least abundant metals: Co^{2+}, Cu^{2+}, Mn^{2+} and Ni^{2+} are shown in Figure 4.13b, where for P1, the concentrations of metals were in following descending order: Mn^{2+}>Ni^{2+}>Cu^{2+}> Co^{2+}. On the other hand, the order of metal concentrations in P2 and P3 was: Ni^{2+}> Cu^{2+}> Mn^{2+}> Co^{2+}. The trends of metal content in these two peaks were very similar, albeit the Ni^{2+} concentration in P3 was significantly higher than P2 by two-fold. The concentrations of Cu^{2+} and Mn^{2+} in P2 however, are significantly more than in P3 by two- and three-fold. The most interesting observation is the concentration of Co^{2+}, which is dangerously near to the detection limit in all samples.

It is indeed interesting to observe that Zn^{2+} was the most abundant metal in P3, and was substantially higher than the amount of Zn^{2+} measured in P2. This could be one of the reasons contributing to the dimer formation, and also may well be related to the slower pace of P3 to reach maximum activity, compared to the other peaks (see section 4.5 i). In this case, Zn^{2+} ions that are possibly occupying the metal sites in P3 may have to be displaced by Co^{2+} in order to be catalytically active. Although the Zn^{2+} concentration in P1 was almost two-fold higher than in P3, it is relevant to note that P1 possesses substantially more Mg^{2+} than P3. The latter metal was measured to be at a higher concentration than Zn^{2+}, and thus may contribute to a faster pace to reach the maximum activity. Mg^{2+} and Zn^{2+} concentrations in P2 however, were inherently similar, and may well reflect the slightly slower pace, compared to P1.

Another main observation in this analysis is that the overall metal content of P1 is substantially higher than the other two peaks, and hence provides a logical explanation for the earlier elution from the ion-exchange column. The metal content of P3 was shown to be higher for Zn^{2+} and Ni^{2+} compared to P2, although the rest of the metals were shown to be in equal amounts or slightly higher in P2. The apparent later elution volume of P2 and P3 from the ion-exchange column could well be the

consequences of lower metal content, where in this case, both samples displayed similar behaviour in most analyses that have been conducted.

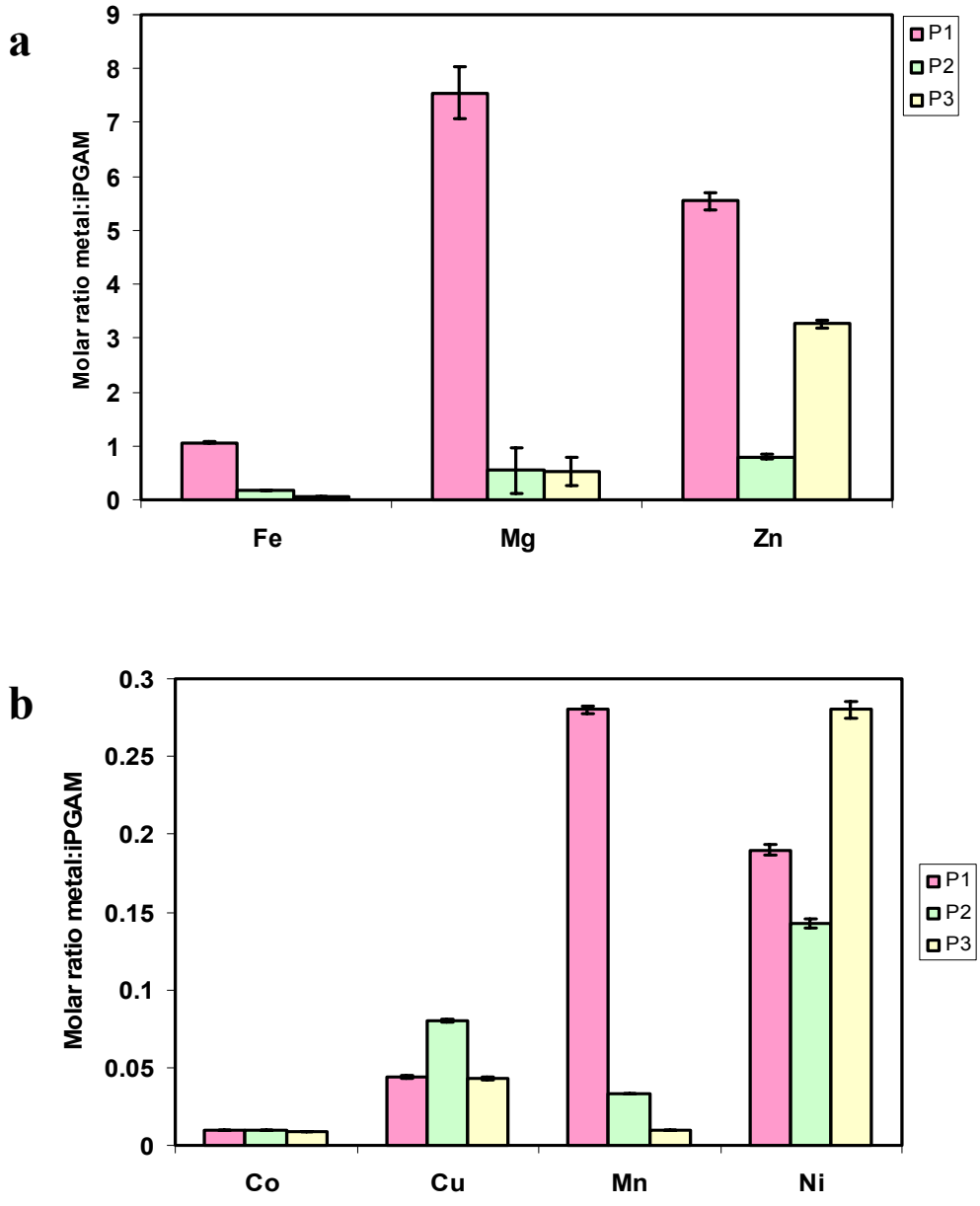


Figure 4.17 ICP-MS analysis of ion-exchange fractions P1, P2 and P3 of *Lm*iPGAM. The relative molar concentrations of metals and purified *Lm*iPGAM are shown. (a) Abundant metals, Fe^{2+}, Mg^{2+}, and Zn^{2+}, with the y axis scale 0 - 8. (b) Low abundance metals, with the y axis scale 0 – 0.25. The following limit of detection values (all μM) was obtained: Fe^{2+}, 0.0313; Mg^{2+}, 0.195; Zn^{2+}, 0.00652; Co^{2+}, 0.000238; Cu^{2+}, 0.000936; Mn^{2+}, 0.00128; Ni^{2+}, 0.000803.

4.9 Conclusion

It was hypothesised in an earlier chapter that the difference in the elution volumes of the three peaks from the ion-exchange column may be due to distinct protein conformational and oligomeric states. Various biochemical and biophysical techniques have been employed in order to investigate whether *Lm*iPGAM does indeed exist in different structural states. The first analysis to be performed was the measurement of the activation of *Lm*iPGAM by Co^{2+}, which showed that in the presence of low Co^{2+} concentration (0.1 mM), the reaction was a slow process, which is possibly due to major conformational change during catalysis. DLS measurements then showed that the trend of hydrodynamic size for *Lm*iPGAM resembled the calculated values, indicating that P1 corresponds to the smaller monomeric closed-form structure, while P2 and P3 most probably correspond to the larger open conformation. The biophysical dimensions however, are slightly different from the crystal values, which may reflect the comparison between proteins in solution and in the crystal lattice. It is noteworthy that from these comparisons, the trypanosomatid and bacterial iPGAMs are divided into two families, which can also be clearly seen from the sequence similarity analysis and the distinct metal-mechanisms for the catalytic function of the two families, apart from the differences in the domain movement of the open-form structures.

Further indication of the existence of multiple species of *Lm*iPGAM was observed in native-PAGE, where multiple bands were observed in P2 and P3, in contrast to the results from SDS-PAGE that exhibited a single prominent band in all the fractions corresponding to the theoretical size of *Lm*iPGAM. Subsequently, the results from gel filtration and SEC-MALS analyses concluded that it was possible to separate the enzyme conformational states by ion-exchange chromatography and by gel filtration. The SEC-MALS analysis, independent of any protein molecular weight standards, also conclusively identified the presence of dimeric iPGAM in both the P2 and P3 samples. Interestingly however, the formation of a dimeric iPGAM in the P1 sample only occurred in the presence of 1 mM Co^{2+}. Nevertheless, the formation of dimer in the presence of metal ions has also been reported in other enzymes, such as *β*-Lactamases from *Pseudomonas aeruginosa* (Danel et al. 2001).

The metal content from ICP-MS indeed showed that the metal content of P1 is mostly dominated by Mg^{2+}, while P3 is strikingly occupied by Zn^{2+}. The presence of both metals in P2 however, was shown to be at equal very low concentrations. These observations could be related to the different rates of activity for all samples after the addition of Co^{2+}, where P3 exhibits a lower pace to reach maximum activity, compared to P1 and P2. Zn^{2+} was shown to possess an inhibitory effect on iPGAM, as discussed later in chapter 6, whereas Mg^{2+} could substantially enhance iPGAM activity, albeit to a lower extent than Co^{2+}. Moreover, the techniques that have been employed in this study indeed confirmed that ion-exchange chromatography possesses the ability to separate *Lm*iPGAM into different conformations and oligomeric states.

Chapter 5:

Hyperactivation of Trypanosomatid iPGAM is Observed *in Vitro*, But Not *in Vivo*

5.1 Aims

Over the years, the regulation of numerous proteins by metal ions has been explored extensively. The requirement of metal ions for the structural and functional integrity of proteins, especially enzymes, is a fascinating aspect of the biochemistry of living cells. Interestingly, over one-third of known proteins are metalloproteins (Ibers and Holm 1980; Tainer et al. 1992; Castagnetto et al. 2002; Page and Di Cera 2006), and such protein-metal complexes tend to favour divalent metals (Page and Di Cera 2006). The alkaline phosphatase superfamily, with trypanosomatid iPGAM as one of its members, has also been typified as requiring divalent metals for its catalytic function. Co^{2+}, as described in chapter 1, plays a vital role in regulating trypanosomatid iPGAM activity, and indeed is the metal that supports the maximum activity of *Lm*iPGAM (Nowicki et al. 2009) and *Tb*iPGAM (Chevalier et al. 2000).

However, looking at the typical metal concentrations in eukaryotic cells (Maret 2010), it is plausible that different scenarios may occur *in vivo*, where the concentrations of Co^{2+} will be much lower than in activity assays and in crystallisation conditions (Nowicki et al. 2009). It is also possible that *in vitro* experimental condition may favour the binding of different metal ions, rather than the native one (Maret 2010).

Taking the previous literature into consideration, this chapter will describe the concentrations of seven transition metal ions in the periodic table (Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}) in the cytosol of *T. brucei*, where iPGAM is located, and also in the purified *Lm*iPGAM. In addition, the specific activities of both the *in vitro* and *in vivo* iPGAMs will be compared, and hyperactivation of the purified enzyme with Co^{2+} will also be discussed in detail. Furthermore, this chapter will discuss the possibilities of whether Mg^{2+} and/or Zn^{2+} play roles as the native metals regulating iPGAM activity *in vivo*.

5.2 Introduction

The majority of metalloproteins have been structurally characterised by analysing ligand preferences and geometries of the divalent metals *in vitro*. It is often difficult however, to determine the correct metal that is involved in regulating the enzyme activity *in vivo*, without understanding the natural environment of the organism from which the protein was isolated (Maret 2010). Over the years, numerous studies have shown that there could be misleading information on the "correct" metal ions thus giving biased understanding of the proteins at the molecular and structural levels. In some cases, non-native metals could be incorporated into the proteins during expression in the culture medium or in the purification process. This phenomenon has been demonstrated for instance, by the incorporation of cobalt instead of zinc in alcohol dehydrogenase when yeast was grown in cobalt (II) enriched medium (Curdel and Iwatsubo 1968). Metal substitution has often been observed to occur *in vitro*, thereby causing hyperactivation with the non-native metals, as exemplified by the incorporation of copper instead of zinc in amino peptidase (Prescott et al. 1985). *In vivo* substitution of divalent metals also has been reported, as in the case of the substitution of zinc by cobalt in carbonic anhydrase in marine diatoms. By contrast however, this particular example demonstrated that the substitution of cobalt is reportedly causing the enzyme to be less active than the native zinc form (Lane and Morel 2000).

Determination of the biologically relevant and correct metals for the catalytic function of a protein normally requires various careful measurements. These include the knowledge of the natural metal bioavailability in the native environment, the suitability of a metal to coordinate in the metal site of a particular protein structure, and also, any significant effects on the activation or loss of protein activities by certain metals. These factors will be discussed in this chapter, where the concentrations of seven biologically relevant metals (Figure 5.1) in the cytosol of *T. brucei* and the purified *Lm*iPGAM, the hyperactivation of *Lm*iPGAM with Co^{2+}, and also the possibilities for other metals to occupy both metal sites in *Lm*iPGAM will be considered. It is noteworthy however, that both metal analyses performed in this study are capable to measure the total metal content in the samples, thus do not

possess the ability to distinguish free metals from any bound metals. Hence, the ICP-OES results for the cytosolic fractions are based on the total (free + bound) metal concentrations in the biological samples. However, since the ICP-MS analysis was performed on the purified *Lm*iPGAM, only metals which are associated with the protein would be quantified, thus provides information on the bound-metal concentrations in the samples.

Periodic Table of the Elements

hydrogen 1 **H** 1.00794																helium 2 **He** 4.002602
lithium 3 **Li** 6.941	beryllium 4 **Be** 9.012182								element name **atomic number** **symbol** atomic weight			boron 5 **B** 10.811	carbon 6 **C** 12.0107	nitrogen 7 **N** 14.00674	oxygen 8 **O** 15.9994	fluorine 9 **F** 18.9984	neon 10 **Ne** 20.1797
sodium 11 **Na** 22.98977	magnesium 12 **Mg** 24.3050											aluminium 13 **Al** 26.981538	silicon 14 **Si** 28.0855	phosphorus 15 **P** 30.97376	sulphur 16 **S** 32.065	chlorine 17 **Cl** 35.453	argon 18 **Ar** 39.984
potassium 19 **K** 39.0983	calcium 20 **Ca** 40.078	scandium 21 **Sc** 44.95591	titanium 22 **Ti** 47.867	vanadium 23 **V** 50.9415	chromium 24 **Cr** 51.9961	manganese 25 **Mn** 54.93805	iron 26 **Fe** 55.845	cobalt 27 **Co** 58.9332	nickel 28 **Ni** 58.6934	copper 29 **Cu** 63.546	zinc 30 **Zn** 65.409	gallium 31 **Ga** 69.723	germanium 32 **Ge** 72.64	arsenic 33 **As** 74.9216	selenium 34 **Se** 78.96	bromine 35 **Br** 79.904	krypton 36 **Kr** 83.798
rubidium 37 **Rb** 85.4678	strontium 38 **Sr** 87.62	yttrium 39 **Y** 88.90585	zirconium 40 **Zr** 91.225	niobium 41 **Nb** 92.90638	molybdenum 42 **Mo** 95.94	technetium 43 **Tc** [98]	ruthenium 44 **Ru** 101.07	rhodium 45 **Rh** 102.9055	palladium 46 **Pd** 106.42	silver 47 **Ag** 107.8682	cadmium 48 **Cd** 112.411	indium 49 **In** 114.818	tin 50 **Sn** 118.710	antimony 51 **Sb** 121.760	tellurium 52 **Te** 127.60	iodine 53 **I** 126.9045	xenon 54 **Xe** 131.293
caesium 55 **Cs** 132.90545	barium 56 **Ba** 137.327	lutetium 71 **Lu** 174.967	hafnium 72 **Hf** 178.49	tantalum 73 **Ta** 180.9479	tungsten 74 **W** 183.84	rhenium 75 **Re** 186.207	osmium 76 **Os** 190.23	iridium 77 **Ir** 192.217	platinum 78 **Pt** 195.078	gold 79 **Au** 196.96655	mercury 80 **Hg** 200.59	thallium 81 **Tl** 204.3833	lead 82 **Pb** 207.2	bismuth 83 **Bi** 208.980	polonium 84 **Po** [209]	astatine 85 **At** [210]	radon 86 **Rn** [222]
francium 87 **Fr** [223]	radium 88 **Ra** [226]	lawrencium 103 **Lr** [262]	rutherfordium 104 **Rf** [261]	dubnium 105 **Db** [262]	seaborgium 106 **Sg** [266]	bohrium 107 **Bh** [264]	hassium 108 **Hs** [269]	meitnerium 109 **Mt** [268]	darmstadtium 110 **Ds** [271]	roentgenium 111 **Rg** [272]	ununbium 112 **Uub** [285]		ununquadium 114 **Uub** [289]				
		lanthanum 57 **La** 138.9055	cerium 58 **Ce** 140.116	praseodymium 59 **Pr** 140.90765	neodymium 60 **Nd** 144.24	promethium 61 **Pm** [145]	samarium 62 **Sm** 150.36	europium 63 **Eu** 151.964	gadolinium 64 **Gd** 157.25	terbium 65 **Tb** 158.9253	dysprosium 66 **Dy** 162.50	holmium 67 **Ho** 164.930	erbium 68 **Er** 167.259	thulium 69 **Tm** 168.934	ytterbium 70 **Yb** 173.04		
		actinium 89 **Ac** [227]	thorium 90 **Th** 232.038	protactinium 91 **Pa** 231.0359	uranium 92 **U** 238.0289	neptunium 93 **Np** [237]	plutonium 94 **Pu** [244]	americium 95 **Am** [243]	curium 96 **Cm** [247]	berkelium 97 **Bk** [247]	californium 98 **Cf** [251]	einsteinium 99 **Es** [252]	fermium 100 **Fm** [257]	mendelevium 101 **Md** [258]	nobelium 102 **No** [259]		

Figure 5.1 Seven biologically relevant divalent metals for trypanosomatid iPGAMs are highlighted

5.3 Low enzymatic activity of iPGAM observed from the cytosolic fractions of *T. brucei*

The first seven enzymes in the glycolytic pathway of trypanosomatids are sequestered in a unique subcellular organelle, the glycosome, with multiple copies in each cell. Interestingly however, iPGAM, together with ENO and PYK is located in the cytosol (Oduro et al. 1980; Chevalier et al. 2000; Guerra et al. 2004). Hence, cytosolic samples were prepared following the method by Misset and Opperdoes 1984, at the Université catholique de Louvain, Brussels, Belgium from both procyclic and bloodstream stages of *T. brucei*. The former stage refers to the form of the parasite when it lives in the midgut of the fly, whereas the latter stage corresponds to the pathogenic form found in mammalian blood.

The integrity of the samples was analysed by SDS PAGE (Figure 5.2), and indicated the absence of proteolytic degradation. S1, S2 and S3 are the procyclic forms, whereas S4 and S7 are the bloodstream forms. The patterns of bands are broadly similar across all the fractions, with the exception of a prominent band with a mass of ~50 kDa in sample S4. The identity of this protein, however, is not known. The concentrations of each sample were determined by Bradford assay, as described in chapter 2. Variations were observed across the samples, from 2.8 to 8.1 mg/mL, hence the results for specific activity and for metal contents were all normalised to 1 mg of protein. The protein concentrations, iPGAM specific activities and estimated iPGAM concentrations of the fractions are as indicated in Table 5.1. No additional Co^{2+} was added into the samples so as to obtain conditions similar to the native catalytic state of iPGAM. It can be seen that the specific activities of iPGAM in the procyclic cytosolic fractions were higher than in the bloodstream-form cytosolic fractions. Similar trends of results were reported by the Brussels research group, where iPGAM activity in both *T. brucei* procyclic and *L. mexicana* promastigote forms are 10 fold higher than the bloodstream form (Paul Michels, unpublished data) (Fuad et al. 2011).

Low specific activities have also been reported in the literature for iPGAM in the absence of added Co^{2+} (Guerra et al. 2004; Albert et al. 2005), and interestingly, this particular enzyme exhibited the lowest activity in the cell lysate, compared to the

other glycolytic enzymes. It is fascinating to observe that trypanosomes expressed their enzymes differentially during different stages of the life-cycle. Although at present, the reason behind this observation is unknown, the results clearly augmented the value of iPGAM as a drug target.

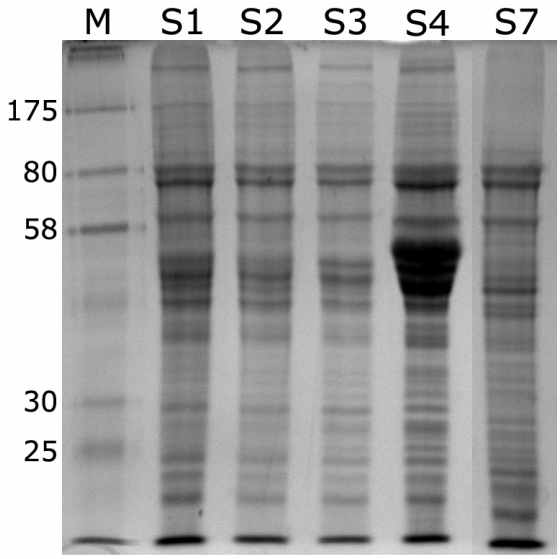


Figure 5.2 SDS-PAGE analysis of the *T.brucei* cytosolic fractions. A 12% SDS-PAGE is shown with 5 µl samples of each cytosolic fraction. The masses (in kDa) of standard marker proteins are indicated in lane M. Abbreviations: M, marker proteins; S1-S3, cytosolic fractions from procyclic *T. brucei*; S4 and S7, cytosolic fractions from bloodstream-form *T. brucei*. The figure is taken from Fuad et al. 2011.

Table 5.1 iPGAM molar concentration in *T. brucei* cytosolic fractions

	S1[a]	S2	S3	S4	S7
protein concentration (mg/ml)	4.0	3.9	2.8	8.1	5.4
specific activity (nmol.min^{-1}.mg^{-1} protein)	16.0	19.7	6.29	1.12	5.80
iPGAM concentration in cytosolic fractions (μmoles/L)[b]	50.1	60.3	13.8	7.12	24.5

[a]Samples: S1, S2 and S3 are procyclic fractions; S4 and S7 are from bloodstream parasites.

[b]The calculation was made using the assumption that pure iPGAM has a specific activity of 21 μmol.min^{-1}.mg^{-1} protein (under the same assay conditions as those used to measure the activity of iPGAM in the cytosol).

5.4 ICP-OES analysis of the cytosol of *T. brucei*

Metal requirements of iPGAMs have been discussed in great detail for trypanosomatids (Chevalier et al. 2000; Guerra et al. 2004, Poonperm 2005) as well as for other organisms (Jedrzejas et al. 2000; Nukui, et al. 2007; Raverdy et al. 2007). However, very little was known about the biologically relevant metal concentrations in the natural environment where the iPGAM was located. It was therefore decided to measure the concentrations of relevant divalent metal ions in cytosolic samples of *T. brucei*, in order to provide or eliminate possible native metal ion candidates *in vivo*.

ICP-OES analysis was employed to measure the concentrations of seven metal ions (Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}) in cytosolic fractions from both procyclic and bloodstream-form *T. brucei*. It is important to keep in mind that it is indeed the total metal concentrations which were obtained by this analysis, where metals bound to macromolecules or to EDTA could not be distinguished from free metal ions. As described in chapter 2, elemental standards were prepared, and gave linear concentration-response plots in both water and in cytosol buffer (see Appendix

I (d)). Any background readings or noise signals collected from the control data were deducted from the readings of the samples. The LOD was determined for each metal ion, as described in the legend for Figure 5.3.

The results from the analysis revealed an interesting trend of metal concentrations, yet provided patterns similar to those observed in other eukaryotic cells (Maret 2010). Mg^{2+}, Zn^{2+} and Fe^{2+} were found to be the most abundant metals (~200, 30 and 15 μM, respectively (Figure 5.3), while Co^{2+}, Cu^{2+}, Mn^{2+} and Ni^{2+} concentrations were determined to be below 1 μM. Interestingly, Co^{2+} was found to be below the limit of detection (<0.035 μM), in which the signals may well be from the background noise of the instrument. It can be seen that the samples from the two different life stages of *T. brucei* gave similar results, which shows that different environmental conditions do not substantially affect the metal content.

The cytosolic concentrations of *Tb*iPGAM may be estimated from the specific activity of the pure enzyme compared to the specific activities within the cytosolic fractions, and give a value of approximately 10-50 μM (Table 5.1). It is clear from the ICP-OES results that the Co^{2+} concentrations in the cytosolic fractions are <0.03 μM, and are therefore insufficient to satisfy even a small proportion of the iPGAM present.

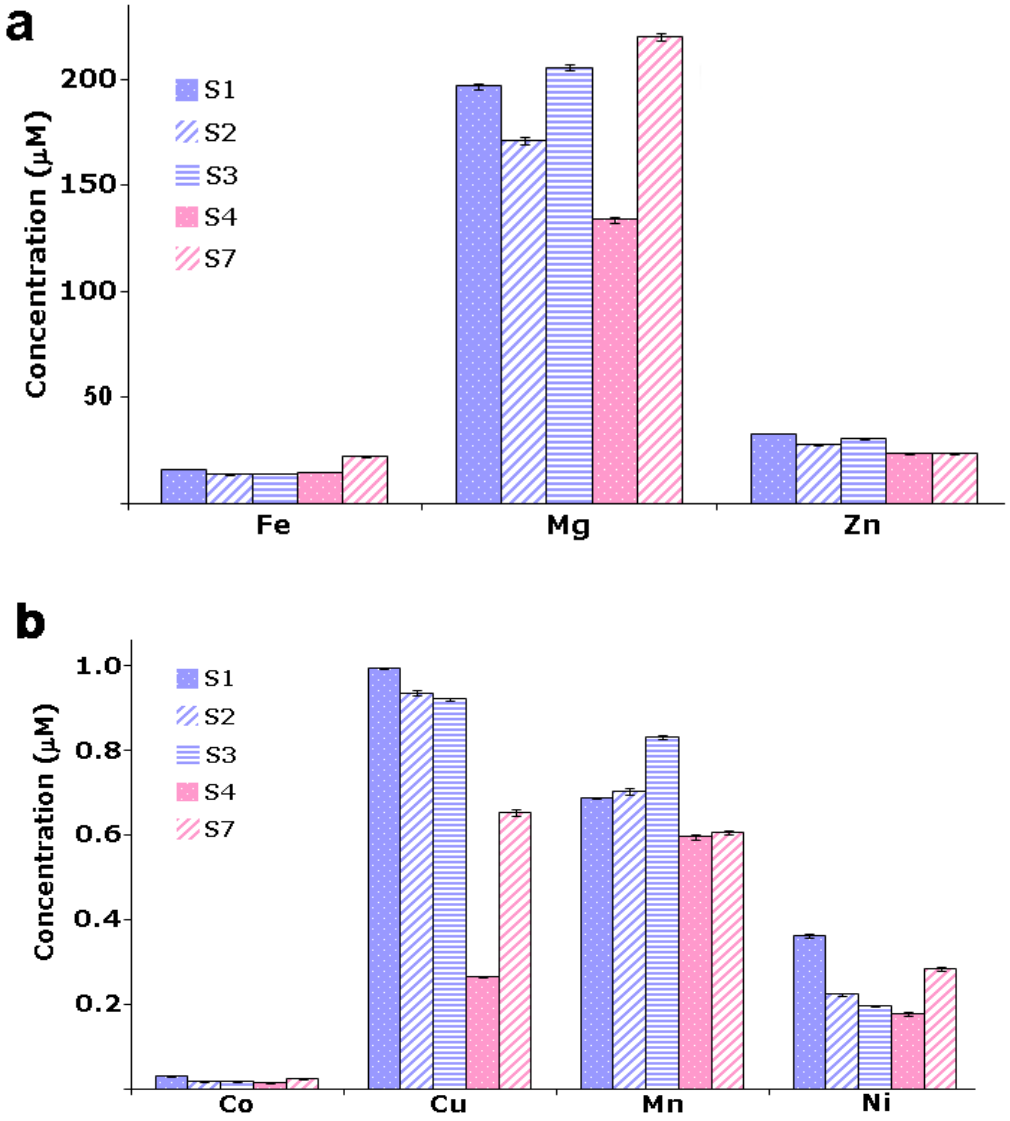


Figure 5.3 ICP-OES of cytosolic fractions from procyclic and bloodstream-form *T. brucei*. The blue bars correspond to the procyclic cytosolic fractions S1, S2 and S3, whereas the pink bars indicate bloodstream cytosolic fractions S4 and S7. (a) Results for the relatively abundant metals, Fe^{2+}, Mg^{2+} and Zn^{2+}, with the y axis scale 0-200 µM. (b) Results for the less abundant metals, Co^{2+}, Cu^{2+}, Mn^{2+} and Ni^{2+}, with the y axis scale 0-1.0 µM. The values for Co^{2+} (0.0130-0.0296 µM) are below the limit of detection (<0.0352 µM). The following limit of detection values (all µM) were obtained: Fe^{2+}, 0.0819; Mg^{2+}, 0.495; Zn^{2+}, 0.122; Co^{2+}, 0.0352; Cu^{2+}, 0.116; Mn^{2+}, 0.0920; Ni^{2+}, 0.104. The figure is taken from Fuad et al. 2011.

5.5 Addition of cobalt *in vitro* can enhance the activity of bacterially expressed *Lm*iPGAM

It is known that the specific activity of *Lm*iPGAM purified in the presence of 10 μM Co^{2+} is ~400 μmol.min^{-1}.mg^{-1} protein (Poonperm 2005). It also has been discussed in chapter 4 that *Lm*iPGAM purified on a Ni^{2+} column required a long period of time (up to 48 hours for P1 and longer for both P2 and P3) to reach its maximal activity after the addition of 0.1 mM Co^{2+}. In this section, similar experiment was conducted on the purified *Lm*iPGAM, but the incubation was performed for one hour, in the presence of higher concentration of Co^{2+} (1 mM). The one hour incubation with 1 mM Co^{2+} was performed based on the results from the previous specific activity measurement (see section 4.5i), which corresponds to the time to reach half of the maximum activity. The initial specific activity for P1, P2 and P3 were calculated to be 19, 18 and 16 μmol.min^{-1}.mg^{-1} protein for the conversion of 3PGA to 2PGA, and when assayed after one hour incubation with 1 mM Co^{2+}, the specific activities of the three samples were increased to 137, 120 and 119 μmol.min^{-1}.mg^{-1} protein, respectively (Figure 5.4). The initial value is similar to the value of 26 μmol.min^{-1}.mg^{-1} protein obtained for pure bacterially expressed *Tb*iPGAM that had also been purified on a Ni^{2+}-IMAC column without the addition of any divalent metals (Collet et al. 2001). This clearly shows that Co^{2+} can strongly stimulate the activity of iPGAM, and indicates that the specific activity *in vivo* would never reach the high specific activity exhibited *in vitro*, taking into account the low concentration of Co^{2+} in the cytosol measured with ICP-OES.

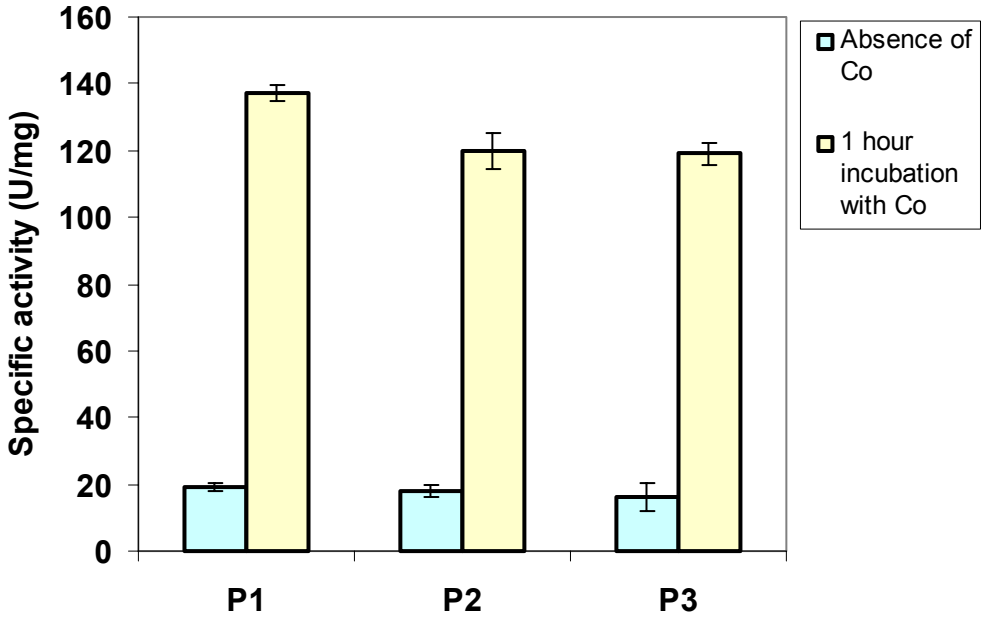


Figure 5.4 The specific activity of the three peaks of *Lm*iPGAM, before and after one hour incubation with 1 mM Co^{2+}. One unit corresponds to the conversion of 1 µmol of substrate min^{-1} mg^{-1} protein under standard conditions.

5.6 ICP-MS analysis of bacterially expressed *Lm*iPGAM

Although the concentrations of metals in the cytosol of the parasite have been determined, it is also important to measure the metal content in the bacterially expressed protein. An important reason for this analysis to be performed is that the sample analysed contains only pure iPGAM, where essentially all other macromolecules must have been washed away during the chromatography steps. Thus, the metal concentrations in the sample would be specific for metals bound only to iPGAM, instead of to other molecules, as in the cytosol. The results would provide a control experiment to the earlier ICP-OES analysis, in order to observe whether the metal concentration trend would be similar to that in the cytosol. In this particular analysis, the metal content of the essentially homogenous purified bacterially expressed *Lm*iPGAM was analysed by the more sensitive ICP-MS technique. Metal

content in the well separated samples from the ion-exchange column: P1, P2 and P3 were analysed, as described in chapter 4. Following similar experimental procedure as ICP-OES, the concentrations of seven different metals (Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}) were measured. Elemental standards were prepared as described in chapter 2, and gave linear concentration-response curves in both water and in enzyme storage buffer (20 mM TEA-HCl, pH 7.6 and 50 mM NaCl) (see Appendix I (e)). Similar to ICP-OES, the LOD has been determined for each metal ion, in both water and the enzyme storage buffer.

In this section, only the metal content of P1 is shown, as a comparison to the ICP-OES results from analysis of the cytosol samples, and exhibited a similar trend of metal content as measured in the cytosol. The metal content of P2 and P3 has been discussed earlier in section 4.8 ii, and showed slight differences to the metal content in P1 and the *T. brucei* cytosolic fractions. Mg^{2+} was identified as the most abundant metal (as it was in the cytosolic fractions), but the concentration of Zn^{2+} was substantially enhanced, and approached the concentration of Mg^{2+} (Figure 5.5). This would indicate that Zn^{2+} is being selectively bound to iPGAM. The high concentrations indicate that iPGAM selectively binds these two metals. The concentration of Co^{2+} was 0.01 μM, above the limit of detection (<0.0002 μM), but very low compared to the concentration of iPGAM (0.99 μM).

Interesting and clear observations have been obtained from both the ICP-MS and ICP-OES analyses. Co^{2+}, despite being the only metal that enables the bacterially expressed *Tb*iPGAM and *Lm*iPGAM to be maximally active, is not present in the enzyme in cytosolic fractions or in enzyme purified in the absence of added Co^{2+}. Two of the most abundant metals, Mg^{2+} and Zn^{2+} exhibited molar ratios of ~8 and ~6 μmol of metal to 1 μmol of *Lm*iPGAM, respectively, indicating the possibility that either or both of these metals are the preferred metal(s) *in vivo*. It is observed from this analysis that the Zn^{2+} concentration has substantially increased and approaches the concentration of Mg^{2+}, which reflects that there is possibility that Zn^{2+} binds specifically to iPGAM during purification. It is also important to note that any metals that are present in *E. coli* cells, where the protein was expressed, may bind to

iPGAM too. Meanwhile, Mn^{2+} that is known to be the preferred metal for bacterial iPGAMs (Kuhn et al. 1995; Jedrzejas et al. 2000a; Jedrzejas et al. 2000b) was present in the purified *Lm*iPGAM at a molar ratio of less than 0.3. It is noteworthy that the concentration of Ni^{2+} was low, despite the purification being done on an IMAC column charged with Ni^{2+}.

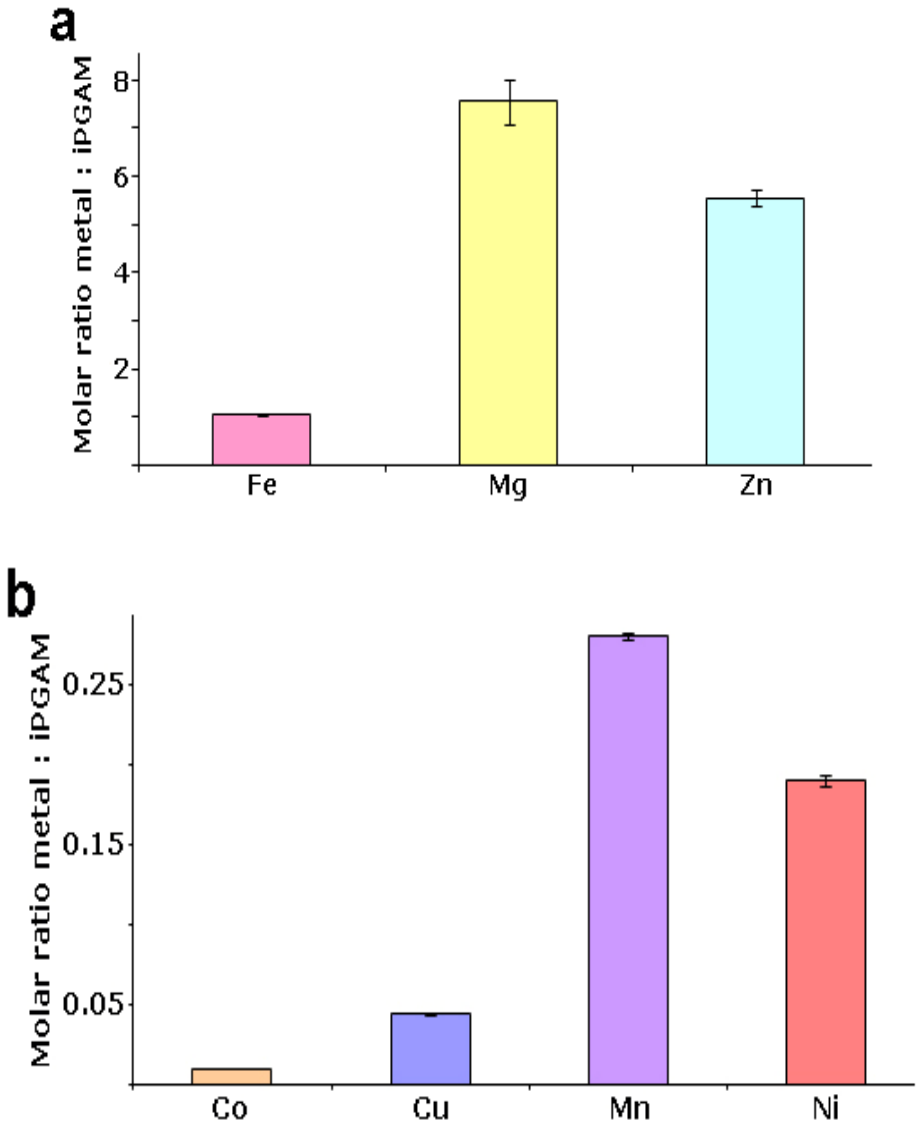


Figure 5.5 ICP-MS analysis of purified *Lm*iPGAM. The relative molar concentrations of metals and purified *Lm*iPGAM are shown. (a) Abundant metals, Fe^{2+}, Mg^{2+} and Zn^{2+}, with the y axis scale 0 - 8. (b) Low abundance metals, with the y axis scale 0 – 0.25. The following limit of detection values (all μM) were obtained: Fe^{2+}, 0.0313; Mg^{2+}, 0.195; Zn^{2+}, 0.00652; Co^{2+}, 0.000238; Cu^{2+}, 0.000936; Mn^{2+}, 0.00128; Ni^{2+}, 0.000803. The figure is taken from Fuad et al. 2011.

5.7 Mg^{2+} and Zn^{2+} could be the native metals required to regulate iPGAM activity *in vivo*

Enzymes are molecules that are present in cells which play important roles in assisting biochemical reactions at the cellular level. It is known that enzymes have often evolved by natural selection to attain a state of optimal catalytic activity, often with adaptation to a particular biological environment. For instance, the activity of the glycolytic enzyme triose phosphate isomerase is limited only by diffusion of substrate onto, and product off, the enzyme, and is therefore thought to have evolved to almost catalytic perfection (Albery and Knowles 1976). There is thus an understandable tendency to correlate the highest specific activity of an enzyme *in vitro* with its 'native' state *in vivo*. This is clearly not the case for trypanosomatid iPGAMs where the maximal activity (by a factor of ten or more) is obtained with Co^{2+}, but the concentration of this metal is essentially undetectable in the parasite cytosol. In this respect, iPGAM is not alone, and carboxypeptidase A (Behnke and Vallee 1972), aminopeptidase (Prescott et al. 1985), alcohol dehydrogenase (Cavaletto et al. 2000), 5'-nucleotidase (McMillen et al. 2003) and catechol dioxygenase (Fielding et al. 2010) are just some among many examples of Zn^{2+}-containing enzymes *in vivo* that are hyperactive when Co^{2+} replaces the Zn^{2+} *in vitro* (Fuad et al. 2011).

The results of ICP-OES indicate that potentially only Mg^{2+} and Zn^{2+} are sufficiently abundant in the parasite cytosol to support the activity of a significant proportion of the iPGAM present (Figure 5.3), whereas Cu^{2+}, Mn^{2+} and Ni^{2+} were all well below the concentration of iPGAM. A similar trend is observed with the ICP-MS analysis of purified *Lm*iPGAM in which only Mg^{2+} and Zn^{2+} are sufficiently abundant to support the activity of the enzyme (Figure 5.5). However, as mentioned earlier, not all metals which are measured in ICP-OES are freely available for iPGAM, although the result is more significant with the ICP-MS analysis. However, the rest of the other metals, especially Co^{2+}, Cu^{2+}, Mn^{2+} and Ni^{2+} can be eliminated from being the native metal supporting iPGAM activity *in vivo*, since the measured concentrations for these metals are insufficient for iPGAM. It thus seems that the biologically relevant metal for this enzyme is likely to be either Mg^{2+} or Zn^{2+}, or

both. As mentioned previously, iPGAM belongs to the metal-dependent alkaline phosphatase superfamily (Galperin et al. 1998) most of whose members require two divalent metals at the active site. Three enzymes are particularly well characterised structurally: *E. coli* alkaline phosphatase (Kim and Wyckoff 1991), *Bs*iPGAM (Jedrzejas et al. 2000; Nukui et al. 2007) and *Lm*iPGAM (Nowicki et al. 2009). Alkaline phosphatase uses two Zn^{2+} ions, *Bs*iPGAM utilises two Mn^{2+} ions, but *Lm*iPGAM by contrast, appears to function with a one-metal mechanism. Of the two metal sites of *Lm*iPGAM, as shown in Figure 1.6 of chapter 1, only site M1 shows high occupancy, with site M2 having only low occupancy (6%) even at 4 mM Co^{2+}. Moreover, when site M2 is occupied, the side chain of Ser75 that is phosphorylated during catalysis adopts a position unfavourable for phospho transfer (Nowicki et al. 2009) (Fuad et al. 2011).

The evaluation of the probability that Mg^{2+} and Zn^{2+} regulate iPGAM activity *in vivo* also requires structural information for the interaction of metals with the neighbouring residues of the metal binding sites. This is important for the identification of possible metals that could be replaced by Co^{2+} when the protein was expressed *in vitro*. Generally, metal cations in proteins prefer to form interactions with donor atoms in the functional groups of certain amino acid side chains, preferably sulphur from cysteine, imidazole nitrogen from histidine, carboxylate oxygen from aspartic or glutamic acid, or oxygen from threonine or serine (Harding et al. 2010). These interactions, which involve a number of atoms within chemical bonding distance (coordination number), would result in the most ideal stereochemistry for each coordination number.

In *Lm*iPGAM specifically, distorted octahedral geometry is observed when Co^{2+} occupying metal site M1, coordinates with the first carboxylate oxygen from Asp425, the second nitrogen from His429, the second nitrogen from His496 and a phospho group oxygen from 2PGA or 3PGA. Apart from these, longer interactions were observed with the second carboxylate oxygen from Asp425 and the bridge oxygen from 2PGA or 3PGA (Figure 5.6a). The second metal site (M2) when occupied with Co^{2+} has tetrahedral geometry with interactions to the following neighbouring atoms: the first carboxylate oxygen from Asp23, oxygen from the

second rotamer of Ser75, first carboxylate oxygen from Asp466 and second nitrogen from His467 (Figure 5.6a) (Nowicki et al. 2009). Based on the interactions mentioned above, it is plausible that the octahedrally coordinated M1 has a preference for Mg^{2+} and Co^{2+}, with coordination number 6 (MESPEUS database, http://mespeus.bch.ed.ac.uk/MESPEUS/_1.jsp). The poorly occupied site M2, on the other hand, has tetrahedral geometry with a preference for Zn^{2+}, and less so for Co^{2+}, with coordination number 4.

The normal coordination number and requirement of donor atoms for Co^{2+} and Mg^{2+} is 6 (Figure 5.7), thus satisfying the preference for regular octahedral geometry. Comparing the possibilities of Mg^{2+} to be in the M1 site instead of Co^{2+}, similar interactions could occur with the neighbouring atoms (Figure 5.6b). However, some possible interactions are rare, for instance, the unusual occurrence of coordination by bidentate carboxylate, or histidine as donor atoms (Harding et al. 2010). This is not impossible however, as the structure of bovine beta-1,4-galactosyltransferase (Ramakrishnan et al. 2004) has illustrated that both interactions could occur when Mg^{2+} occupied the metal sites (http://mespeus.bch.ed.ac.uk/MESPEUS/_1.jsp). Although Ni^{2+} and Mn^{2+} have similar probability to occupy the M1 site with similar coordination geometry in the crystal, the concentrations of the metals as measured by ICP-OES are too small, and hence insufficient to satisfy a 1:1 stoichemistry of metal: protein. Fe^{2+} however, by being the third highest concentration in *vivo* and *in vitro*, has a better chance to occupy the M1 site, as it has very similar requirements for donor atoms, compared to Co^{2+}. Nevertheless, the concentration of Fe^{2+} in the parasite cytosol is only just sufficient to satisfy the binding requirements of iPGAM, even if all of the Fe^{2+} is freely available. By contrast, the total Mg^{2+} content is present in excess concentration and could support the activity of a significant proportion of the iPGAM in the cell. Total Zn^{2+} is also present at a relatively high concentration, and could be able to down-regulate the enzyme activity by occupying site M2, but the extent of this regulation is difficult to evaluate without knowing the concentration of Zn^{2+} freely available. However, it is relevant to note that Zn^{2+} in the purified iPGAM exists at sufficient concentration to be bound to iPGAM, and hence provides a strong

indication of the availability of the metal to regulate iPGAM activity. The effects of these metals on iPGAM activity will be discussed in detail in chapter 6.

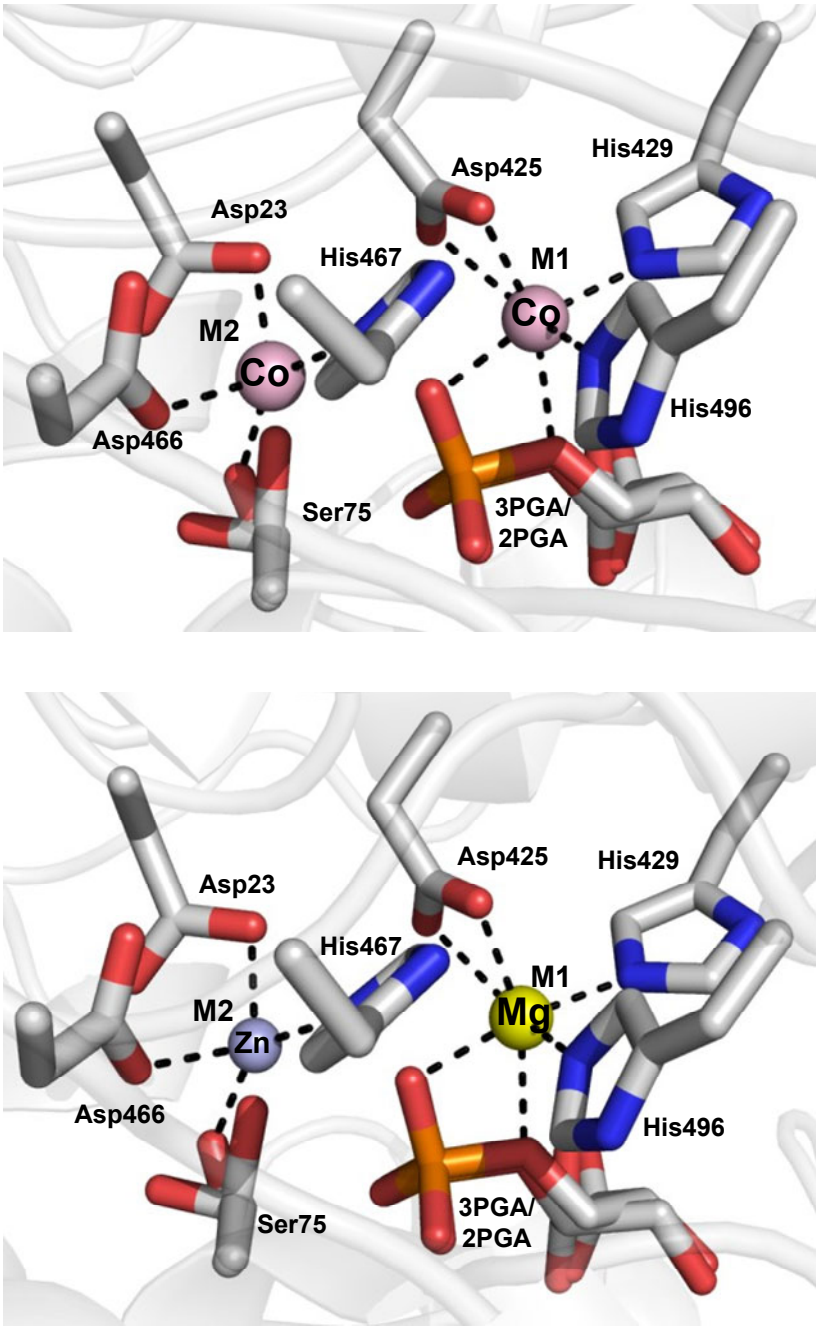


Figure 5.6 Putative metal-binding partners of *Lm*iPGAM. a) Co^{2+} coordinated by neighbouring atoms in both M1 and M2 sites, as observed in the crystal structure of *Lm*iPGAM in the presence of 4 mM Co^{2+}. b) Mg^{2+} and Zn^{2+} modelled into sites M1 and M2, respectively, can readily coordinate the same neighbouring atoms.

Zn^{2+}, compared to most of the other divalent metal ions, is much more promiscuous in its choice of donor atoms and coordination group geometry. A detailed guide of the requirement of donor atoms by Zn^{2+} has been discussed previously (Harding et al. 2010). The usual coordination number for Zn^{2+} is 4 (Figure 5.7), although it has been observed to be coordinated by 5 or 6 donor atoms. In the M2 site of *Lm*iPGAM however, Zn^{2+} is very likely to be tetrahedrally coordinated. Coordination by both monodentate and bidentate carboxylates, and any intermediate patterns are possible with Zn^{2+}, while aspartate, glutamate and histidine are the commonest donor amino acid residues in the catalytic sites. The possible Zn^{2+} interactions with the neighbouring residues in the M2 site are shown in Figure 5.6b.

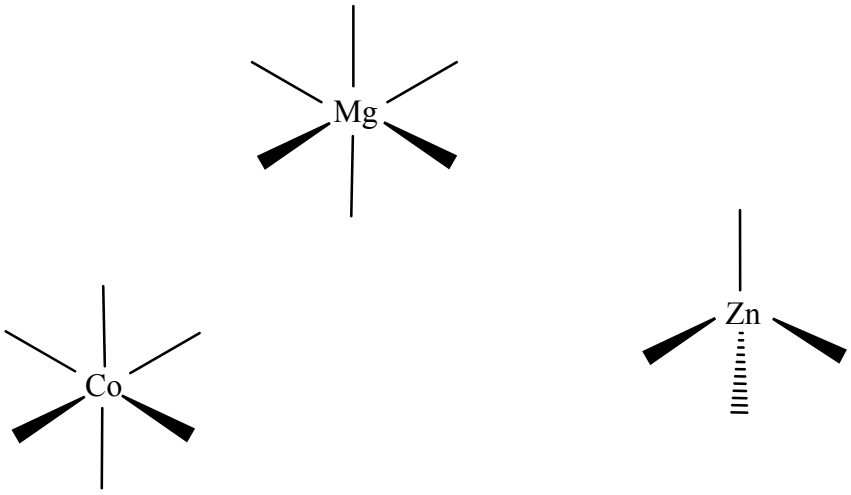


Figure 5.7 The preferred coordination geometries of Co^{2+}, Mg^{2+} (octahedral) and Zn^{2+} (tetrahedral).

It has been suggested that *Lm*iPGAM has a one metal mechanism, instead of two (Nowicki et al. 2009). This is based on the observation in the crystal structure that the side chain of Ser75 can assume three different positions, one of which adopts an unfavourable position for catalysis, when the M2 site is occupied by Co^{2+}. Two hypotheses for the occupation of the metal sites of trypanosomatid iPGAMs by Zn^{2+}

or Mg^{2+} may be proposed: i) Mg^{2+} or Zn^{2+} occupy the M1 site, and site M2 is empty with a one metal mechanism and ii) Mg^{2+} or Zn^{2+} occupies the M1 site, while Zn^{2+} occupies the second site. The geometrical coordination for both metal sites suits the usual preference for both metal atoms, although the interaction of Mg^{2+} with N from histidine, and bidentate carboxylate in M1 is relatively rare, as described previously. It is possible too, that the geometrical coordination might change, as discussed by Poonperm 2005, when both metal sites are occupied by Mn^{2+} or Zn^{2+}.

The potential binding of Mg^{2+} and Zn^{2+} to trypanosomatid iPGAMs was considered further by analysing available published structures of enzymes with Mg^{2+} and Zn^{2+} occupying the catalytic site which form interactions with bidentate carboxylate oxygen from aspartate. This is one of the important interactions to consider in the M1 site, as Co^{2+} interacts with both carboxylate oxygens from Asp425. The MESPEUS database (http://mespeus.bch.ed.ac.uk/MESPEUS/_1.jsp) was used for this purpose, where the metal-bidentate carboxylate interactions were searched for both Mg^{2+} and Zn^{2+}. For Mg^{2+}, 137 hits of PDB structures with resolution 1.6 Å or better were obtained for the interactions with bidentate carboxylate oxygen. However, there are some difficulties in bidentate definitions, as there are three possibilities of interactions: i) both O atoms from one carboxylate group bind to one metal ion: ii) one O atom from the carboxylate group binds to two metal ions iii) both O atoms from one carboxylate group bind to two different metals (Dr. Marjorie Harding, personal communication). In the case of possible interactions in the M1 site, only the first possibility was taken into account, where five pairs of O atoms (ten altogether) were obtained from the initial 137 hits. Zn^{2+}, on the other hand, exhibited more common interactions with bidentate carboxylates, exhibiting 23 pairs of O atoms (46 altogether) out of 73 initial hits. It should be noted that the coordination number for the search was fixed at six, assuming that similar geometrical coordination was obtained when Mg^{2+} or Zn^{2+} occupy the M1 site.

These results suggest that both Mg^{2+} and Zn^{2+} have the ability to bind to the two O atoms of the carboxylate from Asp425, although the interaction was found rare for Mg^{2+}, compared to Zn^{2+}. Meanwhile, it is also plausible that Mg^{2+} forms an interaction to monodentate carboxylate from Asp425, with an additional interaction

with an O atom from a water molecule. Dudev and Lim 2004 have described that in such cases, water molecules have a crucial role in stabilising the monodentate carboxylate-binding mode of water-rich Mg^{2+} complexes. Nevertheless, the key factor in determining the carboxylate-binding mode is the availability and type of interactions between the metal-free carboxylate O atom and its neighbours (Dudev and Lim 2004). In the case of Zn^{2+} however, as mentioned above, the latter metal prefers coordination number 4 for the interaction, thus the coordination geometry might change too if the M1 site is occupied by Zn^{2+}.

The metal-donor atom distances observed in protein crystal structures are 2.07±0.10 Å for the interaction between Mg^{2+}-O from monodentate carboxylate, while the value for Zn^{2+}-O is 1.99±0.05 Å (Harding et al. 2010). Although these values are estimated for the interaction of metals with monodentate carboxylate, the metal-distances for Mg^{2+} and Zn^{2+}-bidentate carboxylate also fall within the same range (Figure 5.8). The mean value for O-metal distances is 2.14±0.16 Å and 2.20±0.17 Å for Mg^{2+} and Zn^{2+}, respectively.

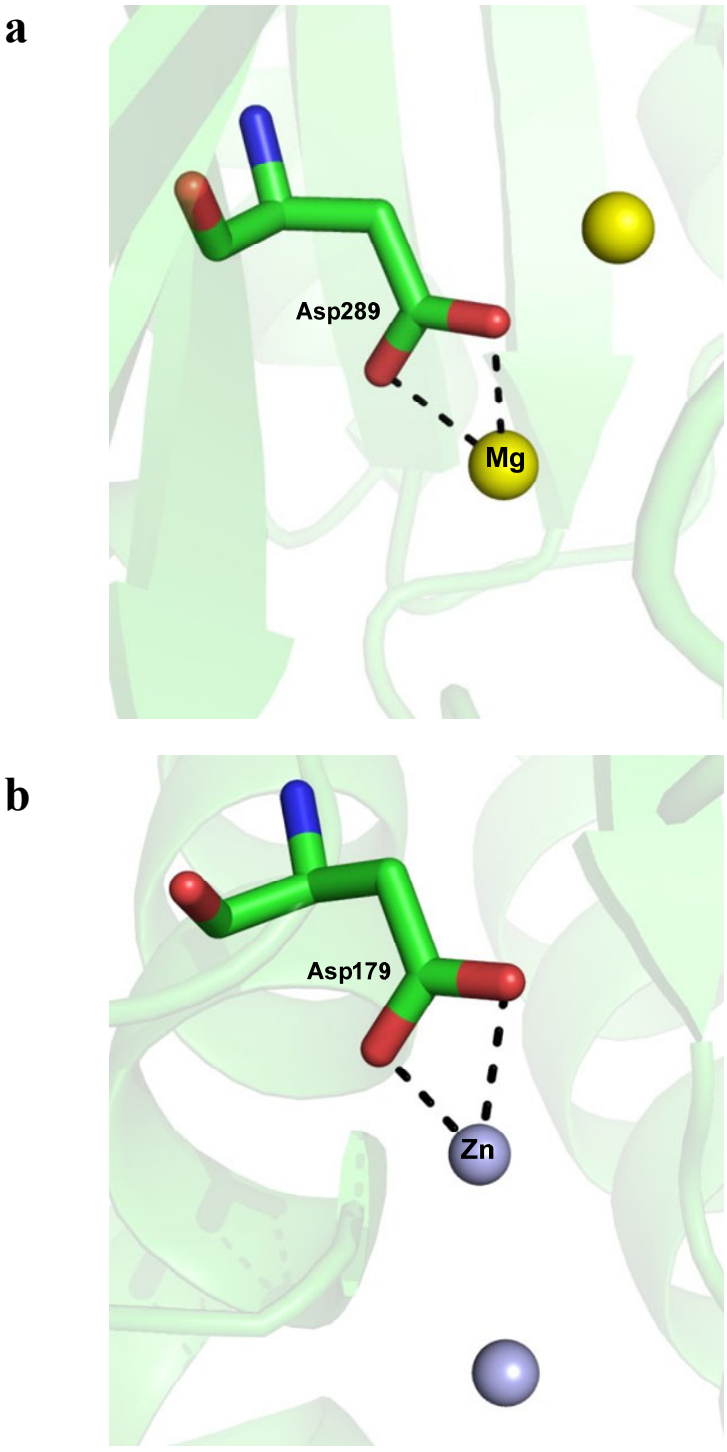


Figure 5.8 Possible metal-bidentate carboxylate coordination. a) Examples of Mg^{2+}-bidentate carboxylate coordination (inositol 1,3,4-trisphosphate 5/6-kinase, PDB ID 1Z2O) b) Examples of Zn^{2+}- bidentate carboxylate coordination (aminopeptidase from *Aeromonas proteolytica*, PDB ID 1LOK).

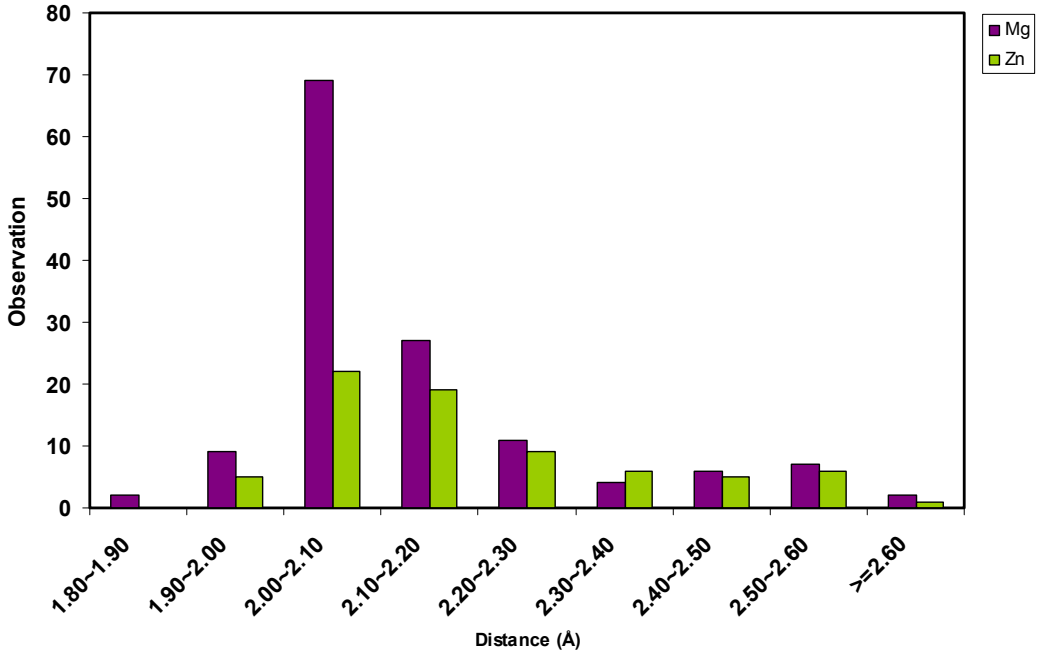


Figure 5.9 Distribution for observed distances for metal-bidentate carboxylate oxygen in Mg^{2+} and Zn^{2+}. The observation corresponds to the number of proteins showing the various metal-bidentate carboxylate distances. The data were obtained from the MESPEUS database.

Possible effects on the catalytic mechanism are anticipated with changes in the interactions of the neighbouring atoms with the different metals. Poonperm (2005) reported that the substitution of Mn^{2+} and Zn^{2+} into both metal sites by crystal soaking led to differences in metal coordination geometries and protein-ligand binding at the active site of *Lm*iPGAM, which caused the activity of iPGAM to be decreased. In the presence of 1 mM Mn^{2+} and 0.01 mM Co^{2+}, the coordination geometry of both the M1 and M2 sites changed to trigonal bipyramidal. Zn^{2+} on the other hand, when present at 0.5 mM concentration together with 0.01 mM Co^{2+}, exhibited a tetrahedral coordination in M1, while forming five tight interactions in M2, causing the geometry to change to trigonal bipyramidal. While Mn^{2+} prefers the latter coordination, this is not the case with Zn^{2+}, as it is observed to be less preferred compared to four-coordination. Nevertheless, both findings observed that Ser75, which becomes phosphorylated during catalysis, had moved towards the phospho group of the phosphoglycerate, thereby being in a favourable position to form a phosphoserine intermediate, and explaining the inhibitory effect of both metals. It

was suggested by this study that Mn^{2+}- and Zn^{2+}-bound proteins represent iPGAM trapped in a non-productive intermediate stage during catalytic activity, in contrast to Co^{2+}-bound iPGAM, which represents an early or final stage of the catalytic reaction. In another experiment, in the presence of a higher concentration of Ni^{2+} (4 mM) and similar concentration of Co^{2+} (0.01 mM), no significant changes could be observed with lower occupancy of the metal in M2 site. This observation indicated that Ni^{2+} was not required, and had a low preference for metal coordination in the active site of iPGAM (Poonperm 2005).

The major problem in observing the effects of these metals on the activity of iPGAM is the intolerance of the coupling enzymes in the activity assay to higher concentrations of metals. Hence, an assay system which can be carried out in the presence of different divalent metal ions with minimal interference on the coupling enzymes was developed, as described in chapter 6.

The loss of activity, with metal substitution in the metal sites has been reported in other enzymes. Human carbonic anhydrase II is one of the enzymes that display negligible activity when the native metal Zn^{2+}, which was tetrahedrally coordinated by three histidine residues and a water molecule was replaced by other divalent trace metals: Cu^{2+}, Mn^{2+} and Ni^{2+}. The coordination number around the metal was increased, to penta-coordinated with Cu^{2+}, and hexa-coordinated with both Mn^{2+} and Ni^{2+} (Håkansson and Wehnert 1992). Interestingly, the tetrahedral geometry was restored when Co^{2+} occupied the metal site. Nevertheless, the reaction of Co^{2+}-substituted enzyme was slowed down, thus proposing that Zn^{2+}, with its natural abundance and coordination geometry, is the native metal that supports enzyme activity *in vivo*. This might be the case for *Lm*iPGAM too, where Co^{2+} conversely, activated the enzyme much better than the native metals *in vivo*. It is not possible at this stage to decide definitively which metal is the biologically relevant one, and indeed it may be that both metals can support trypanosomatid iPGAM activity. Hence, it is crucial to measure iPGAM activity in the present of these significant metal ions, to indicate which metals may support or inhibit the enzyme activity. These experiments will be discussed in the following chapter.

5.8 Conclusion

The findings in this chapter have culminated in a better understanding of the hyper activation of trypanosomatid iPGAM with Co^{2+}. The bacterially expressed protein purified with a Ni^{2+} column would never reach the maximum activity observed with samples purified in the presence of Co^{2+}, unless additional Co^{2+} was added into the samples. A comparison of the specific activities for both samples remarkably showed an almost 20-fold difference, indicating that Co^{2+}, unlike other metals, could activate iPGAM to a much higher extent. This observation however, is only anticipated with samples purified *in vitro*, since Co^{2+} concentrations *in vivo* were found to be much lower than in Co^{2+}-purified samples and in the crystal structure.

Through ICP-MS and ICP-OES analysis, it is clear that Co^{2+} concentrations are much lower than the other biologically relevant metal ions. Both the bacterially expressed protein and the cytosol of the parasites exhibited low concentrations of Co^{2+}, which indicates that it is not sufficient to support iPGAM to a maximal activity. By comparing the concentrations of *Tb*iPGAM in the cytosolic fractions (~10-50 μM), to Co^{2+} concentrations *in vivo* which are less than 0.03 μM, it is clear that the stoichiometry is insufficient to satisfy even a small proportion of the iPGAM present. A similar observation was reported by ICP-MS for the bacterially expressed protein, where the concentration of Co^{2+} was only 0.01 μM. Although the value is above the limit of detection (<0.0002 μM), it is still exceptionally low compared to the concentration of the bacterially expressed iPGAM in the same sample (0.99 μM).

Mg^{2+} and Zn^{2+} were found to be at higher concentrations compared to the other divalent metal ions. Both metals exist at concentrations that are more than sufficient to satisfy a significant proportion of iPGAM either *in vivo* or *in vitro*. However, as mentioned earlier, the concentrations of metals that are freely available to bind to iPGAM may be overestimated, since these analyses do not distinguish free metals from bound ones. Nevertheless, the results from ICP-MS analysis provide strong indications for this hypothesis, since the sample analysed is purified iPGAM. In this respect, any metals present, in theory should be specifically bound to iPGAM. Moreover, the low activity observed from the bacterially expressed protein (~21

U/mg) may also be supported by Mg^{2+} and/or Zn^{2+}. The M1 and M2 metal sites in iPGAM exhibited structural preferences for Mg^{2+} and Zn^{2+}, respectively, based on coordination geometry and the possibilities of interactions with the neighbouring atoms. In the next chapter, further analysis on the activities of iPGAM in the presence of these two significant metal ions, and also with the other relevant divalent metal ions will be discussed.

Chapter 6:

The Development of an Assay System for Trypanosomatid iPGAM in The Presence of Metal Ions

6.1 Aims

The previous chapter has discussed metal concentrations in the cytosol of *T. brucei*, and also in the bacterially expressed protein from *L. mexicana* in great detail. There are similarities in the trends of metal concentrations in both samples, where Mg^{2+}, Zn^{2+} and Fe^{2+} are in greatest abundance, compared to the least abundant metals; Mn^{2+}, Ni^{2+}, Cu^{2+} and Co^{2+}. In particular, Co^{2+} is present below the limit of detection, which suggested that *in vivo* cellular conditions lacking Co^{2+} are unable to support the maximal activity of iPGAM, and Mg^{2+} and/or Zn^{2+} may assume this role. The significance of these findings required further analysis on the activities of iPGAM in the presence of various metal ions, which could be carried out initially on the *in vitro* conditions from the bacterially expressed protein in the laboratory. As detailed in chapter 2, the assay system used to measure the activity of iPGAM is based on the oxidation of NADH, by coupling iPGAM with ENO, PYK and LDH. The well known intolerance of enolase towards divalent metal ions (Lee and Nowak 1992; Poonperm 2005), and the requirement of Mg^{2+} by both enolase and PYK (Ainsworth and Macfarlane 1975; Brewer 1981) provide challenges in assaying the activity of iPGAM in the presence of a series of divalent metals. Hence, the aim for this chapter is to develop an assay system which could be carried out at minimal metal interference to the latter coupling enzymes using a more modern, fast and cost saving multimode plate reader technique, and hence, could provide indications of the significant roles of these divalent metals to iPGAM activity.

6.2 Introduction

In 1988, Johnson and Price had reported that iPGAM activity could be monitored using an assay system called the 'coupled-quenched' procedure, or also known as 'stop and sample' assay. In this experiment, a portion of iPGAM and the substrate 3PGA was incubated together before disruption of the enzyme activity. A small amount of the enzyme catalysed reaction which contained a mixture of 3PGA and 2PGA was added into a mixture containing the rest of the enzyme components, including the other three coupling enzymes. The end absorbance reading after absorbance decrease at A_{340} nm was recorded and compared to calibration curves to obtain the amount of 2PGA produced in the initial reaction. The rationale behind the development of the assay is to avoid the interference of a range of divalent metals to be tested on iPGAM, with the rest of the coupling enzymes in the whole system. This type of assay corresponds to a discontinuous assay, where iPGAM as a reversible enzyme acted to convert 3PGA to 2PGA and vice versa before the next reaction in the presence of the latter coupling enzymes takes place. By contrast, the standard continuous assay is a non-reversible way of monitoring the enzyme activity, purely because the presence of the rest of the coupling enzymes includes the essentially irreversible step catalysed by PYK. Figure 6.1 showed the differences between the two types of assays.

Our novel assay utilises the recent advancement in micro plate reader format, which exhibits dissimilarities from the traditional cuvette approach described in Johnson and Price (1988), and also from the continuous assay system that has routinely been implemented in the laboratory. Considering the volume in the cuvette system is 10 fold higher than the micro plate wells, the consumption of the reagent in the latter instrument could be brought down which is more sufficient to carry out large scale analysis. The micro plate with 96 wells also made it feasible to perform multiple samples for an analysis in a much shorter period of time.

This chapter will describe the recent progress in optimising the micro plate reader technique to measure the absorbance of NADH at A_{340} nm, which is directly involved in measuring iPGAM activity. Subsequently, the chapter will discuss the

details of the discontinuous assay which was implemented in the newly developed multimode plate reader technique. The effects of a series of divalent metals will also be described in this section.

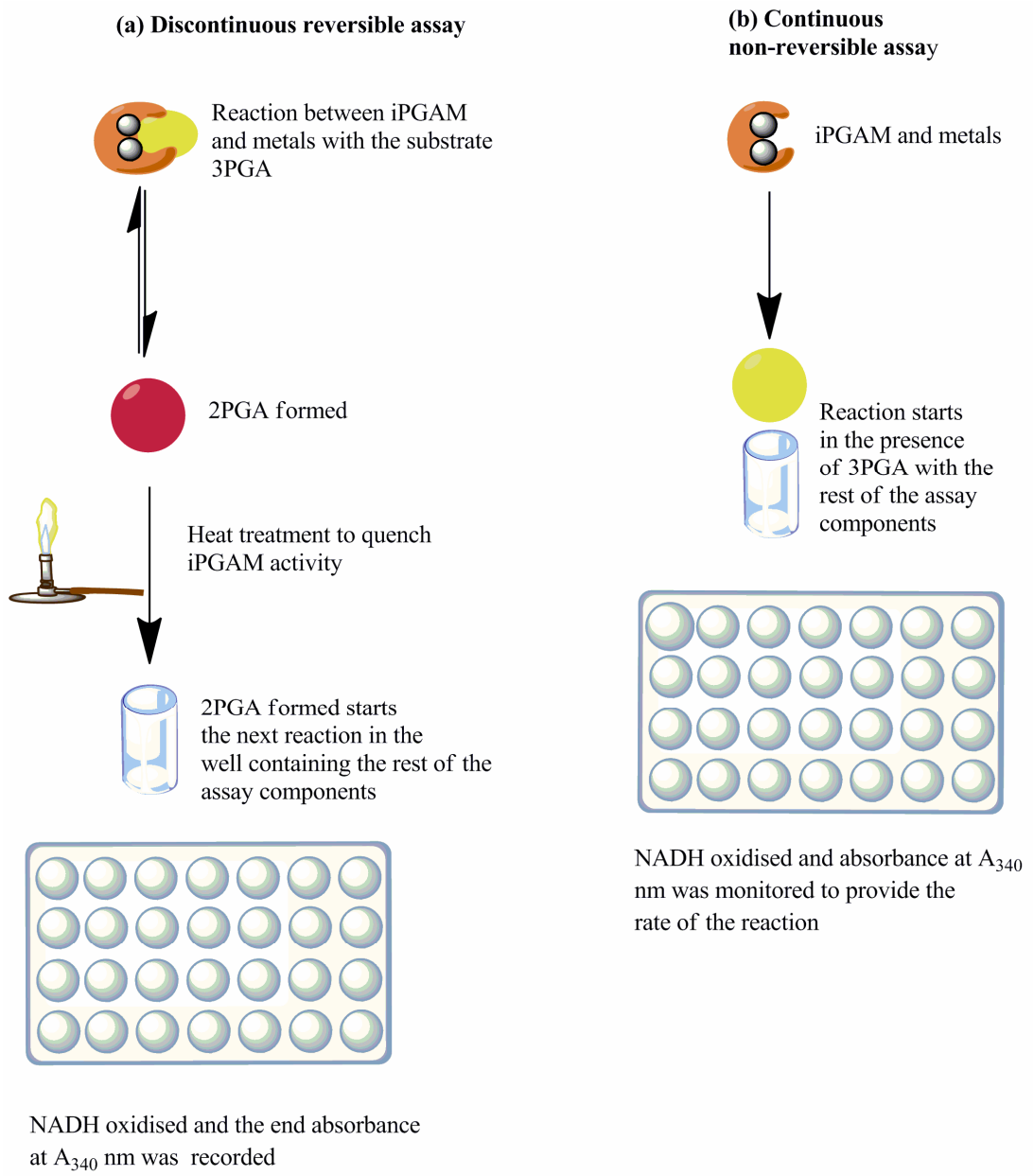


Figure 6.1 The steps in the discontinuous and continuous assays for measuring iPGAM activity.

6.3 Optimisation of the plate reader format for assaying phosphoglycerate mutase enzymatic activity

a) Determination of the parameters for the assay components.

i) NADH concentration

The standard coupling enzyme assay was monitored by the oxidation of NADH to NAD^+, which occurred during the conversion of pyruvate to lactate by lactate dehydrogenase. In this section, a stock concentration of 6.4 mM NADH was diluted two-fold to get a series of different concentrations: 1.6 mM, 0.8 mM, 0.4 mM, 0.2 mM, and 0.1 mM. The rest of the enzyme components were specified at the standard concentrations in the normal continuous assay, as described in chapter 2. The starting absorbance was recorded and plotted in a line graph, as shown in Figure 6.2.

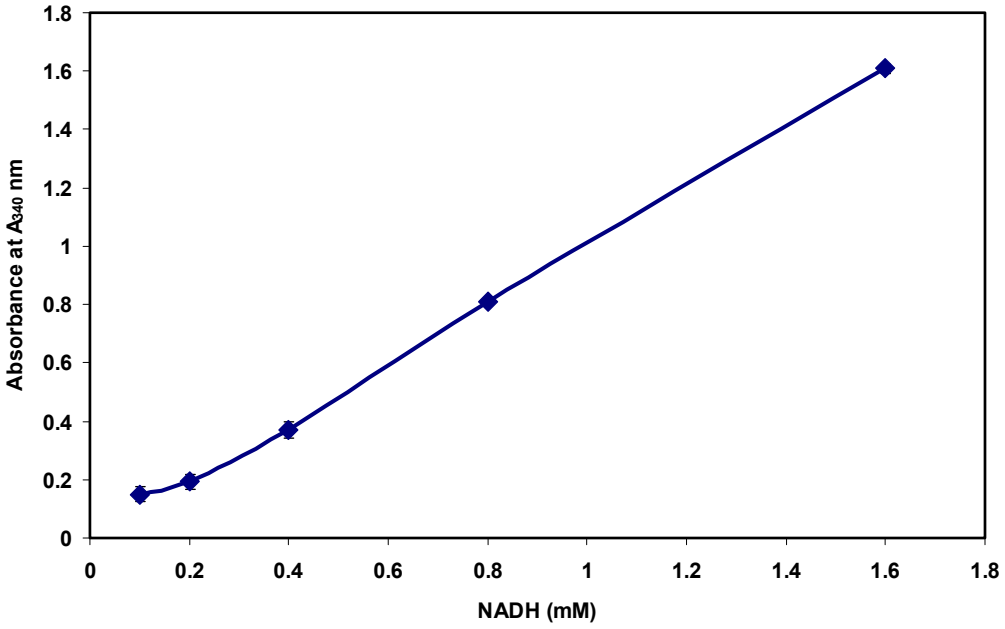


Figure 6.2 Determination of the optimal NADH concentration in the *Lm*iPGAM assay mix. 0.8 mM to 1.6 mM NADH concentrations provided sufficient signal for the analysis. Lower NADH concentrations exhibited inadequate signal as monitored in the spectrophotometer. Errors were calculated but the values are too small to show up.

The results indicated that NADH concentrations from 0.8 mM to 1.6 mM are sufficient to obtain good starting absorbance for the reaction. On the other hand, lower concentrations of NADH (0.4, 0.2 and 0.1 mM) did not provide adequate signal for the starting absorbance. It was therefore decided that future experiments would be employed in the presence of 0.8-1.6 mM NADH.

ii) Enzyme concentration

Different enzyme concentrations were tested to obtain the optimal velocity in determining the rate of the reaction. A range of protein concentration was tested: 0.00125, 0.0025, 0.005, 0.01, 0.02 and 0.04 mg/ml as the final concentrations in the well. The results, as depicted in Figure 6.3, demonstrated a linear correlation between the amount of enzyme used, and the velocity. In this initial experiment, the reaction was conducted for 30 min, which means that the enzyme concentration required would be low. It was later determined that an enzyme concentration in the range of 0.0025 to 0.005 mg/ml was found to be sufficient for the multimode plate reader assay.

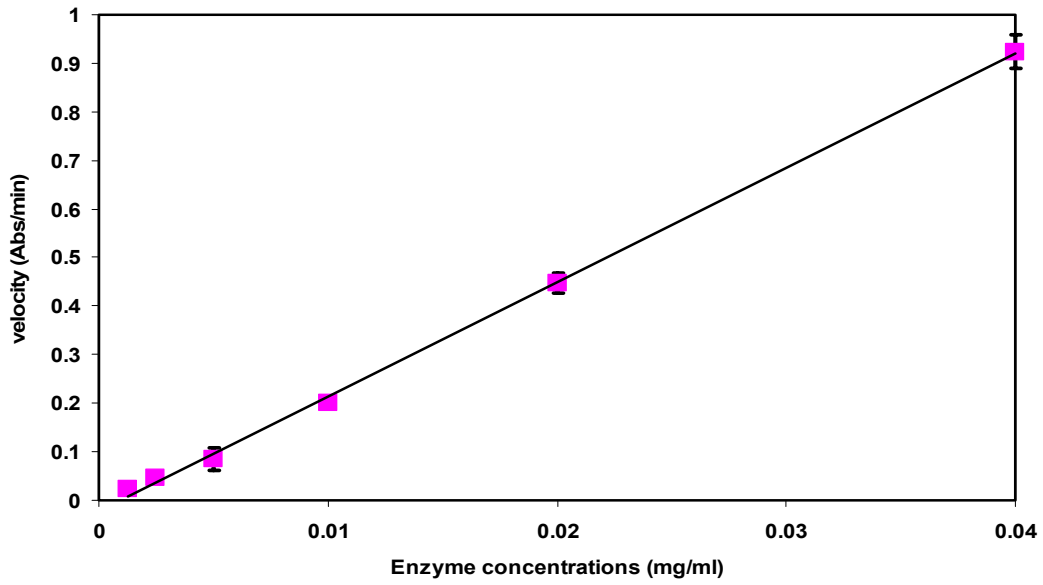


Figure 6.3 Determination of the *Lmi*PGAM concentration required for the multimode plate reader assay. Errors were calculated but the values are too small to show up.

b) Specific activity determination of *Lm*iPGAM and *Tb*iPGAM

In chapter 4, the difference in the specific activities of *Lm*iPGAM was explained in detail, with regard to the different forms and oligomerisation status of each peak eluted from the ion-exchange column. In the optimisation steps for the multimode plate reader however, determination of the specific activity was carried out only by using P1 from the ion-exchange step for *Lm*iPGAM and purified protein from the gel filtration step for *Tb*iPGAM. Both samples were treated with 1 mM Co^{2+} for 45 min at 25°C, and samples without added Co^{2+} were used as controls. The specific activity of *Lm*iPGAM without the presence of Co^{2+} was determined to be 21 U/mg, and increased to 183 U/mg after incubation with Co^{2+} for 45 min. *Tb*iPGAM provided similar observation with 33 U/mg of specific activities before the addition of Co^{2+}, and 190 U/mg after incubation for 45 min (Figure 6.4).

As anticipated, these two set of results are very similar to the results obtained using the standard spectrophotometric cuvette assay (chapter 4), and also suggested that *Tb*iPGAM, as *Lm*iPGAM, is hyperactivated by Co^{2+} at standard conditions. These results also supported the previous findings in chapter 5, where Co^{2+}, when present at very low concentrations in the cytosol of *T. brucei*, could not sustain the activity of iPGAM at a high level. Alternatively, the activity may have been maintained at a low level by the presence of different metal ions. It is also noteworthy that the specific activity before the addition of Co^{2+} in *Tb*iPGAM (33U/mg) is similar to the value obtained by Collet (2001), which is 26 U/mg. The enzyme was also purified with a Ni^{2+} column.

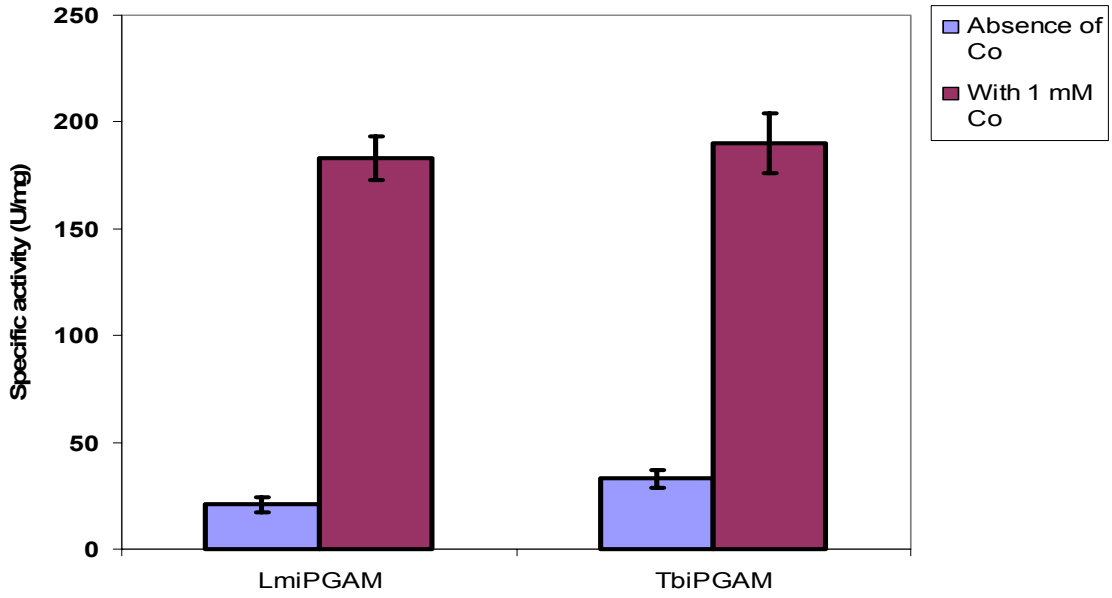


Figure 6.4 Specific activities of *Lmi*PGAM and *Tbi*PGAM before and after the incubation with 1 mM Co^{2+} for 45 min. One unit corresponds to the conversion of 1 µmol of substrate min^{-1} mg^{-1} protein under standard conditions.

i) Kinetic values: V_{max} and K_m values for *Lmi*PGAM and *Tbi*PGAM

Following the determination of the specific activities for both *Lmi*PGAM and *Tbi*PGAM, the kinetic values V_{max} and K_m were determined. Varying concentrations of substrate 3PGA was added into the reaction mixture at a final concentrations of 1.5 mM, 0.75 mM, 0.375 mM, 0.1875 mM, 0.09 mM, 0.05 mM and 0.02 mM. The K_m values for *Lmi*PGAM were determined for the three peaks from ion-exchange chromatography, and from the gel filtration column for *Tbi*PGAM. To note, 1 mM Co^{2+} is present in all samples prior to the kinetic measurement.

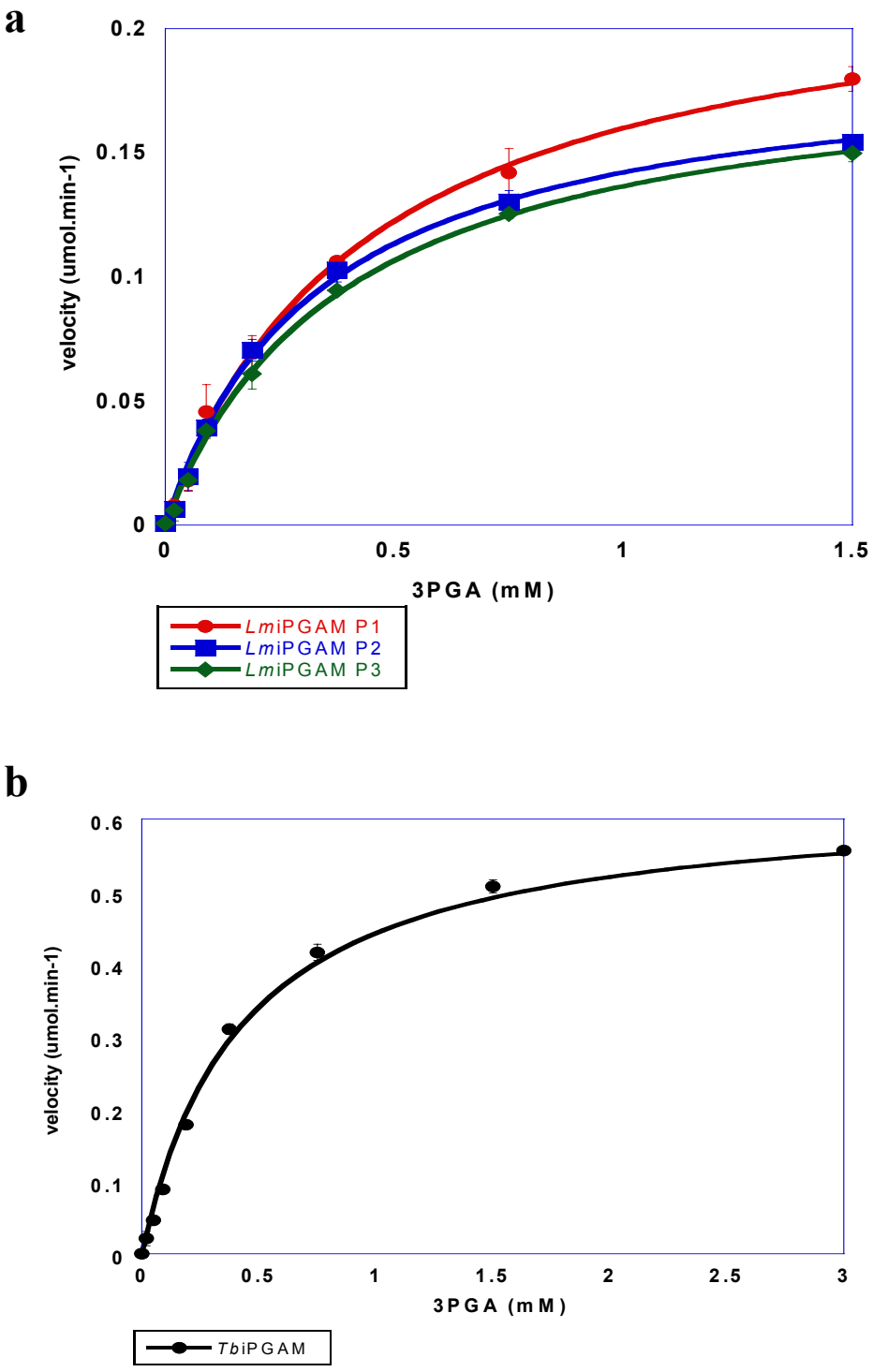


Figure 6.5 Michaelis Menten plot for the determination of the K_m values for 3PGA in a) *Lm*iPGAM and b) *Tb*iPGAM. The plots were generated using the program Kaleidagraph. Errors were calculated but the values are too small to show up.

The K_m values for 3PGA for the three peaks from ion-exchange column are tabulated in Table 6.1, and similar to each other, with 0.45 ± 0.03 mM for P1, 0.33 ± 0. 03 mM for P2, and 0.39 ± 0.03 mM for P3 (Figure 6.5a). For *Tb*iPGAM, the K_m value was determined to be 0.43 ± 0.04 mM (Figure 6.5b). Nevertheless, these values were higher than *Lm*iPGAM that has been purified by a Co^{2+}-IMAC column, which were 0.18 ± 0.01 mM (Poonperm 2005) and 0.27 ± 0.02 mM (Guerra et al. 2004) (Table 6.1). Meanwhile, the K_m values that have been reported for Co^{2+}-TALON purified *Tb*iPGAM and EDTA-treated *Tb*iPGAM purified by Ni^{2+}-NTA agarose column after reactivation with Co^{2+} were 0.15 ± 0.02 mM (Chevalier et al. 2000) and 0.1 mM (Collet et al. 2001), respectively. Tighter bindings of 3PGA to the active site in the present of higher concentrations of Co^{2+} in Chevalier et al. 2000; Poonperm 2005 and Guerra et al. 2004, or in the absence of other metals reported in Collet et al. 2001 was expected since the specific activities and the V_{max} values are higher with enzyme that has been purified in the presence of Co^{2+}. Since this study is focusing on the activity using the forward reaction, the K_m value for 2PGA is not determined. The observation of the hyperactivity of trypanosomatid iPGAM with Co^{2+} and how other metals could support the enzyme activity at lower levels will be discussed in detail in section 6.4.

Table 6.1 The *K*m values for *Lm*iPGAM and *Tb*iPGAM from this study, and the previously reported values.

Proteins	***K*m values (mM) from this study**	***K*m values (mM) from other study**
*Lm*iPGAM	P1 0.45 ± 0.03 P2 0.33 ± 0.03 P3 0.39 ± 0.03	0.27 ± 0.02 (Guerra et al. 2004) 0.18 ± 0.01 (Poonperm 2005)
*Tb*iPGAM	0.43 ± 0.04	0.15 ± 0.02 (Chevalier et al. 2000) 0.1 (Collet 2002)

c) Metal tolerances of the coupling enzymes: lactate dehydrogenase, pyruvate kinase and enolase.

i) Tolerance of LDH to divalent metal ions

In order to observe the effects of metals on iPGAM, the tolerances of the coupling enzymes to divalent metals have to be determined. These results are important to draw borderlines of the amount of metals which could be tested on iPGAM, without significant interference to the other coupling enzymes. Lactate dehydrogenase is the last enzyme in the coupling enzyme assay, converting pyruvate, the ultimate product in glycolysis, to lactate. In the assay mixture, LDH from rabbit muscle was utilised for the reaction. Since this is the first trial that was employed to observe the tolerances of the coupling enzyme towards divalent metal ions, only Co^{2+} was included in the experiment. Hence, using pyruvate as the substrate without the presence of the other coupling enzymes, the activity of LDH was measured in the presence of Co^{2+} ion. Concentrations of Co^{2+} ranging from 0, 0.1, 1, 1.5, 2, 3, 4, 5 and 10 mM were tested to observe their effects to LDH.

The results (Figure 6.6) concluded that the activity of LDH is decreased by Co^{2+}, with the remaining activity ~60% at higher concentrations of Co^{2+}. Mg^{2+} however, was always present in the assay buffer at 5 mM concentration. It was therefore decided that the next enzyme, which is PYK, would be tested together with LDH, in the absence of added Mg^{2+} in the assay buffer.

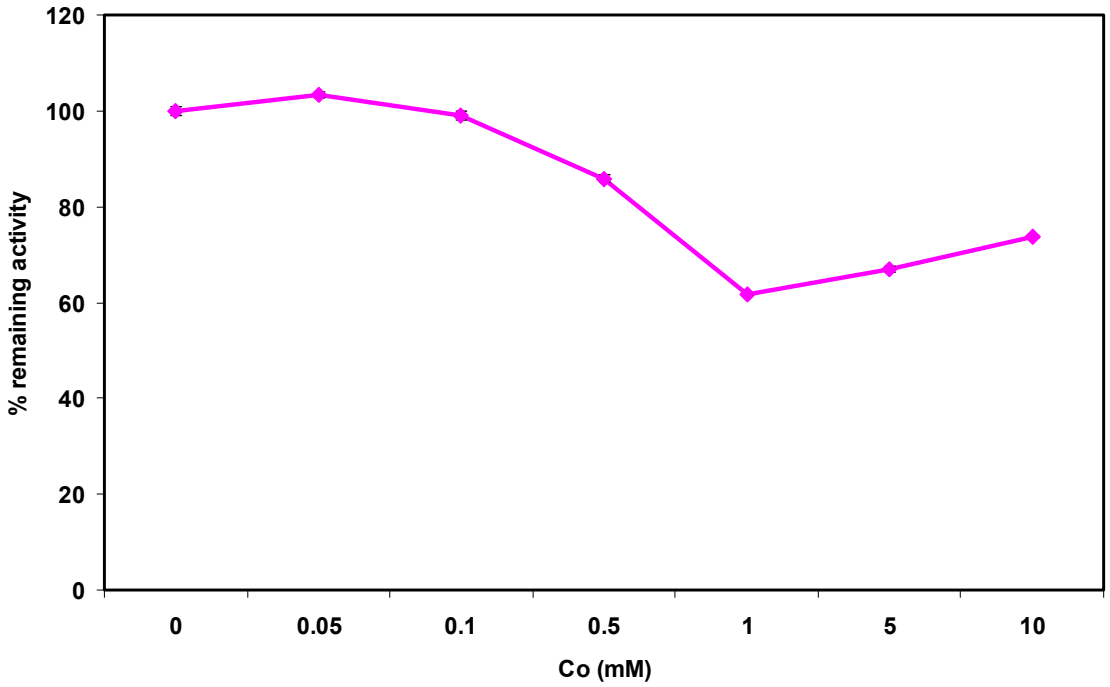


Figure 6.6 Effect of Co^{2+} on the activity of LDH. The reaction was started with pyruvate as the substrate, in the presence of 5 mM Mg^{2+}. Errors were calculated but the values are too small to show up.

ii) Tolerances of PYK and LDH to divalent metal ions

PYK catalyses the conversion of phosphoenolpyruvate (PEP) to pyruvate, and plays a role as the final enzyme in glycolysis. In this study, the activity of PYK from rabbit muscle was tested in the presence of Mg^{2+}, Mn^{2+}, Fe^{2+}, Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}, in concentrations ranging from 0, 0.1, 1, 1.5, 2, 3, 4 and 5 mM. It is important to note that the initial activity was assayed in the absence of any divalent metals, and for the purpose of testing the effects of Mg^{2+}, this metal was omitted from the assay mixture. Activation of rabbit muscle PYK by Co^{2+}, Mn^{2+} and Ni^{2+} had been reported previously (Mildvan and Cohn 1965; Bygrave 1966; Ainsworth and Macfarlane 1975). Meanwhile, it is well-known that rabbit muscle pyruvate kinase requires Mg^{2+} and a monovalent cation, preferably K^{+} for the reaction between phosphoenolpyruvate and ADP is to occur (Kachmar and Boyer 1953; Ainsworth and Macfarlane 1975). The initial reaction (in the absence of metal ions) exhibited

almost no activities of the enzymes in the absence of the divalent metal ions. It is clearly shown in Figure 6.7 that Co^{2+} and Mn^{2+} activated the enzymes, even up to two-fold more strongly than Mg^{2+}. The rest of the metals (Cu^{2+}, Fe^{2+}, Ni^{2+} and Zn^{2+}) however do not support PYK activity.

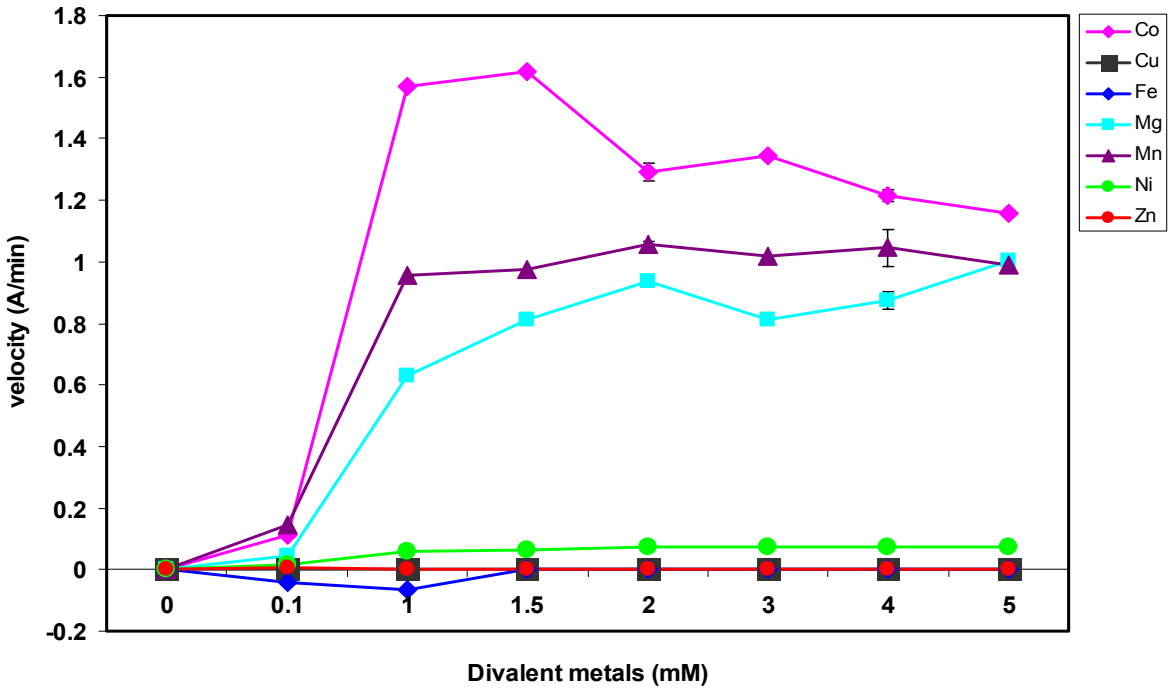


Figure 6.7 Effects of various divalent metals on the activities of PYK and LDH. The reaction was started with phosphoenolpyruvate as the substrate, in the absence of Mg^{2+} in the assay buffer.Errors were calculated but some of the values are too small to show up.

iii) Tolerances of ENO, PYK and LDH to divalent metal ions

In the glycolytic pathway, ENO utilises 2PGA as substrate, producing PEP as the ninth and penultimate metabolite in the pathway. The poor tolerance of yeast ENO towards a number of divalent metals such as Mn^{2+} and Zn^{2+} has been discussed previously (Lee and Nowak 1992). These observations however, contradict older reports by Parisi and Vallee (1969) and Brewer (1981) which reported that transition metal ions such as Mg^{2+}, Mn^{2+}, Ni^{2+}, Co^{2+} and Zn^{2+} are activators of yeast ENO. In the present study, two sets of analysis were conducted, where in the first experiment

the activity of the coupling enzymes was assayed in the presence of Mg^{2+} in the assay buffer (Figure 6.8a). For the purpose of assaying the effects of Mg^{2+} to the enolase, the second experiment was performed in the absence of the metal. The experimental condition employed was the standard condition for PYK, except that the activity was started by 2PGA, which is the substrate of ENO.

In the first experiment (Figure 6.8a), significant depletion of the enzyme activity was observed as the concentration of Zn^{2+} and Ni^{2+} increased. The enzymes exhibited better tolerance to Co^{2+}, although the activity gradually decreased as the concentration of the metal rose. Mn^{2+} however, maintained the enzyme activity almost at 100% remaining activity at the whole range of concentrations. With Zn^{2+}, the 100% inhibitory effect was observed as the concentration increased to 0.1 mM, while the enzymes exhibited better tolerance to Ni^{2+}, with 100% inhibitory effect at 5 mM.

The second experiment (Figure 6.8b) was conducted only in the presence of Co^{2+}, Mn^{2+} and Mg^{2+}. The initial activity was nearly zero when Mg^{2+} was omitted from the assay mixture, exhibiting similar observations as the previous experiment where PEP started the reaction in the absence of ENO. It is clear that Mn^{2+} and Mg^{2+} are able to support the activities of the coupling enzymes, with Mg^{2+} being the more active. Co^{2+}, on the other hand, exhibited a smaller activating effect than Mg^{2+} or Mn^{2+}.

The preliminary test for assaying the tolerances of the coupling enzymes towards the divalent metals concluded that *Lm*iPGAM activity has to be assayed in a condition where the presence of metals concentrations would have none or minimal effects towards the coupling enzymes. Considering the fact that any significant observations would be masked by the effects of metals towards PYK or ENO, either as inhibitor or activator, iPGAM activity has to be determined prior to the addition of the other coupling enzymes. Therefore, in the next section, the development of the discontinuous assay, or the coupled quenched procedure will be discussed in detail.

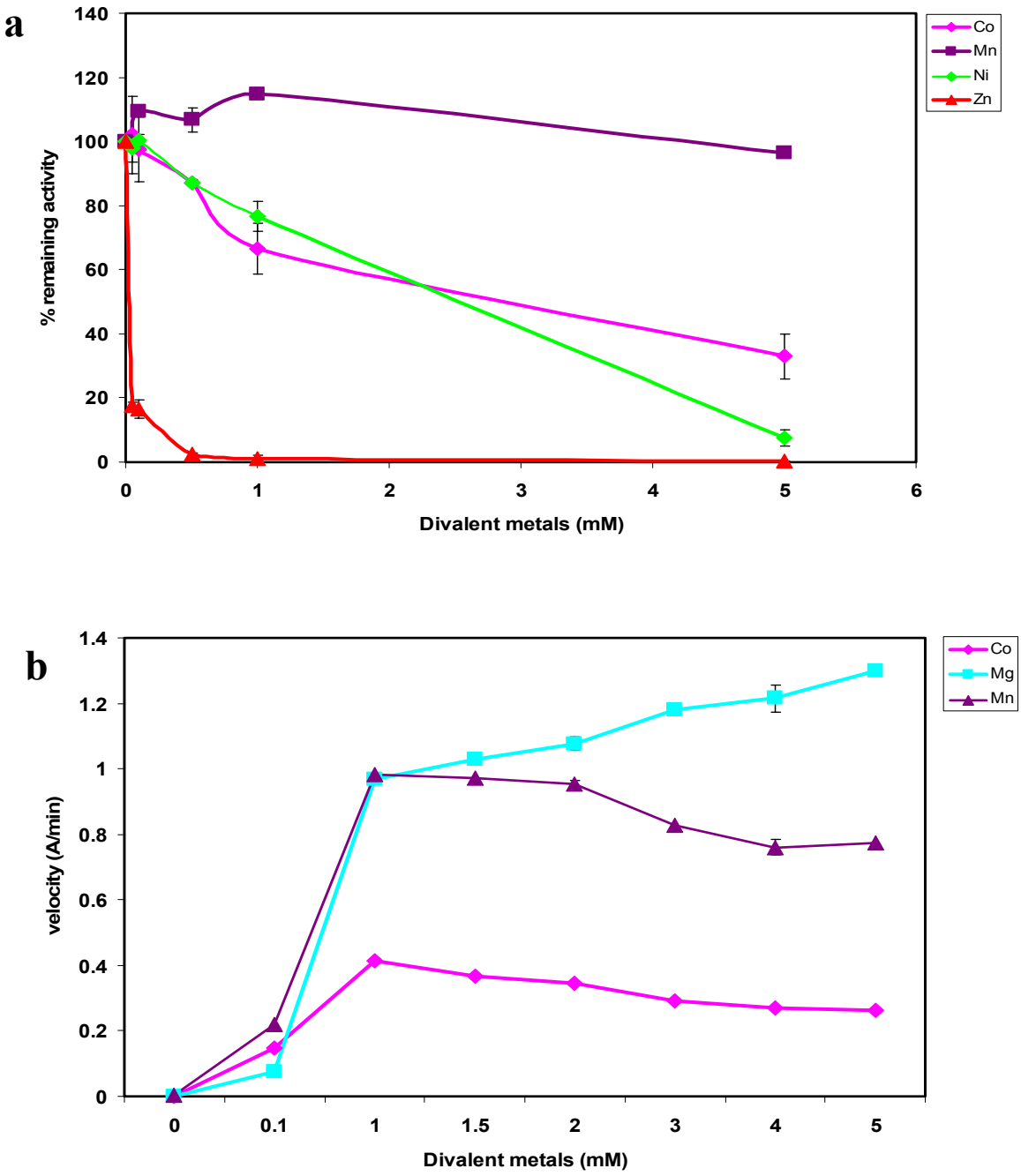


Figure 6.8 a) Effects of Co^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+} on the activities of the coupling enzymes (enolase, PYK and LDH). The reaction was started with 2PGA as the substrate in the presence of 5 mM Mg^{2+} b) A separate experiment of effects of Co^{2+}, Mg^{2+} and Mn^{2+} on enolase, PYK and LDH. This analysis was done in the absence of Mg^{2+} in the assay buffer. Errors were calculated but some of the values are too small to show up.

6.4 The development of the multimode-plate reader discontinuous assay

A discontinuous assay measurement is one of the best methods of testing iPGAM activity in the presence of metal ions, without interference of metals on the activity of coupling enzymes. Apart from minimising the poor tolerance of ENO towards Zn^{2+}, the assay may limit the possibility of "false" activation of the enzyme by Mg^{2+}, by separating iPGAM activity from the rest of the coupling enzymes. These advantages however, were challenged by the limitations of conducting the end-point assay in which only small amounts of 2PGA are present. Moreover, iPGAM is inherently difficult to assay because neither the substrate nor product can be readily measured directly. Hence, the activity measurements required not only iPGAM, but the other three coupling enzymes: ENO, PYK and LDH to obtain convenient spectroscopic signals in which changes due to iPGAM activity could be observed. Traditionally, using the non-reversible continuous assay, the product (2PGA) would be taken away by the next enzyme (ENO) and thus favour the forward direction. In the discontinuous assay however, the scenario is different. iPGAM, in the absence of the other coupling enzymes, not only would convert 3PGA to 2PGA, but would catalyse the reverse reaction, until the equilibrium between 3PGA and 2PGA is attained. The ratio of 3PGA to 2PGA at equilibrium was not known prior to these measurements, but was initially assumed to be of the order of 1 to 1. It was later found to be 1 to 0.08 (see section 6.5a). This situation is often the major difficulty in optimising this type of assay, in which the length of incubation for iPGAM and 3PGA, together with the amount of enzyme and substrate used become the major challenge in obtaining sufficient product for measurement. These are all factors that have to be considered to determine the time to reach equilibrium. Figure 6.9 illustrates the reversible and non-reversible iPGAM reactions.

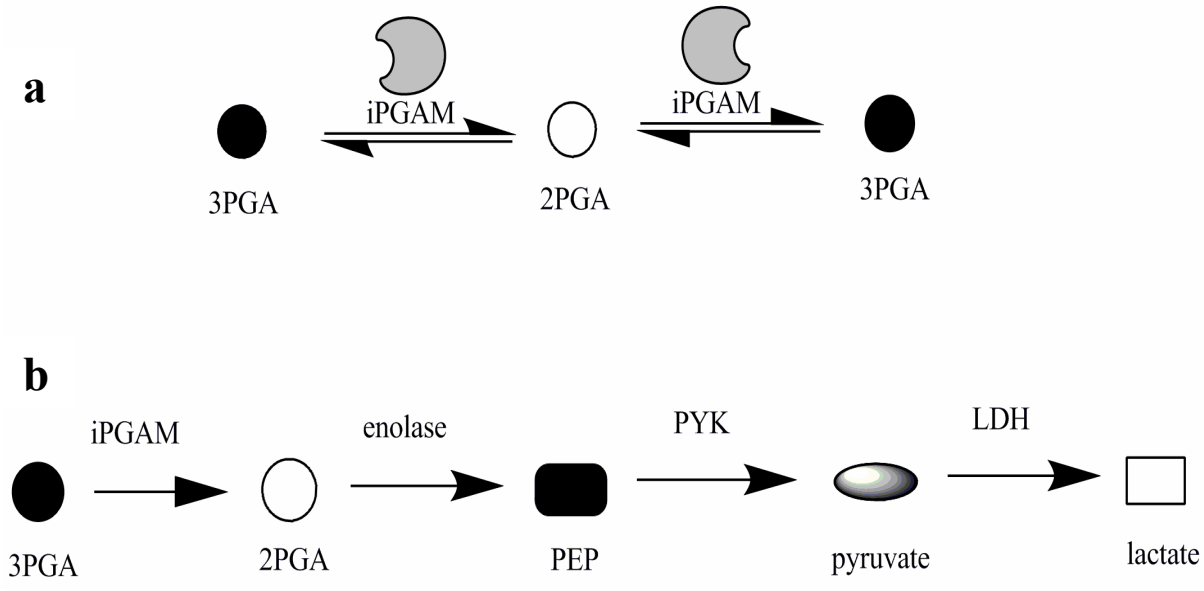


Figure 6.9 The reactions catalysed by iPGAM. a) Reversible reaction in the presence of iPGAM alone. b) Essentially non-reversible reaction in the presence of iPGAM, along with the rest of the coupling enzymes: ENO, PYK and LDH.

The principle of the discontinuous assay method involves the reaction of iPGAM with its substrate 3PGA in the presence of divalent metal ions in the first tube, before being deactivated (for example by rapid heating, changes in pH or in some conditions by adding trichloroacetic acid that denatures and precipitates almost all proteins (Price and Nairn 2009)). The 2PGA which is present in the first tube is the substrate used to start the next reaction containing the rest of the coupling enzymes and the assay components, preferably in the plate reader wells. The reaction was monitored by the decreased absorbance of NADH at A_{340} nm, and the concentration of 2PGA was determined by comparing the end absorbance of NADH to the standard curves, which were generated with known concentrations of 2PGA. This concentration of 2PGA would correspond to 1/20 of the concentration in tube 1 because of the 20-fold dilution from tube 1 to the plate reader well. The discontinuous assay method proved to be convenient for monitoring the activity of iPGAM in the presence of metals that affect the activities of ENO, PYK and/or LDH.

a) The generation of the standard curves in the presence of various metal ions

As mentioned above, iPGAM activity is assessed by the concentration of 2PGA in the first tube. Hence, the calibration standards are vital to ensure the validity of the concentration of 2PGA, which reflects the activity of iPGAM. In this particular step of assay development, 2PGA at known concentrations were prepared (0.63 mM, 0.31 mM, 0.16 mM, 0.08 mM and 0 mM) and added to the reaction mix in the well of the plate reader containing the rest of the enzyme components (chapter 2), excluding iPGAM and 3PGA. Prior to the analysis, the 2PGA had been heated at 100°C for 2 min and showed that the metabolite is stable to boiling, reflecting the suitability of the method of quenching in this analysis.

A further step in assay development concerned the tolerance of the coupling enzymes to divalent metals at different concentrations, bearing in mind the 20-fold dilution from tube 1 to the plate reader well. Thus, three different concentrations of metals (0.05, 0.1 and 0.5 mM) for each Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+} were added separately to individual reactions (in the presence of the usual 5 mM Mg^{2+} and 50 mM K). Although Ca^{2+} was not included previously in the ICP-OES or ICP-MS analysis, the presence of this metal in the cytosol of *T. cruzi* (Docampo et al. 1993; Scott et al. 1997) and *T. brucei* (Xiong and Ruben 1998) made it an interesting candidate to test, along with the other metal ions. Taking into account the 20-fold dilutions made up upon adding 5 µl from the original reaction mix into 100 µl total volume of final reaction, the effects of up to 1, 5 and 10 mM metals could be tested on iPGAM activity in tube 1. Altogether, 25 calibration standards were constructed, at three different concentrations for all eight metals, with a control in the absence of added metal ions, except for 5 mM Mg^{2+} and 50 mM K^{+} in the assay buffer (see Appendix I (f)). Many of the standard curves were generated using linear fit; $y=mx + c$, where m is the slope and c is the y-intercept, and some others were plotted using a polynomial regression with the formula $y=ax^2 – bx + c$. The R^2 values for these curves were >0.99, indicating the goodness of the fit.

As observed in the standard curves (see Figure 6.10 and Appendix I (f)), no significant effects on the activity of the coupling enzymes could be observed even at

the highest concentration tested, which was 0.5 mM for some metals, such as Ca^{2+}, Co^{2+} (Figure 6.10a), Mg^{2+}, Mn^{2+} and Ni^{2+}. Fe^{2+} started to show a decrease in the activity of the coupling enzymes at 0.5 mM, albeit no significant effects could be seen at 0.05 and 0.1 mM (Figure 6.10b). Zn^{2+} and Cu^{2+} were the two metals which showed the greatest effect on enzyme activity. Decreased activity in the presence of these metal ions could be observed with increasing metal concentrations, starting from the lowest concentrations: 0.05 mM and up to 0.1 mM and 0.5 mM, and hence, resulting in a non-linear regression (see Appendix I (f)).Considering these results, and for the purpose of mimicking the *in vivo* condition as in the cells, while keeping the concentration biologically relevant, the highest concentration that would be tested in tube 1 containing iPGAM, metals and 3PGA was standardised to be 1 mM. Hence, all the end absorbance values for the activities of iPGAM in the presence of specific metal ions could be compared to standard curves in the presence of 0.05 mM metals, taking into account the 20-fold dilutions made after quenching.

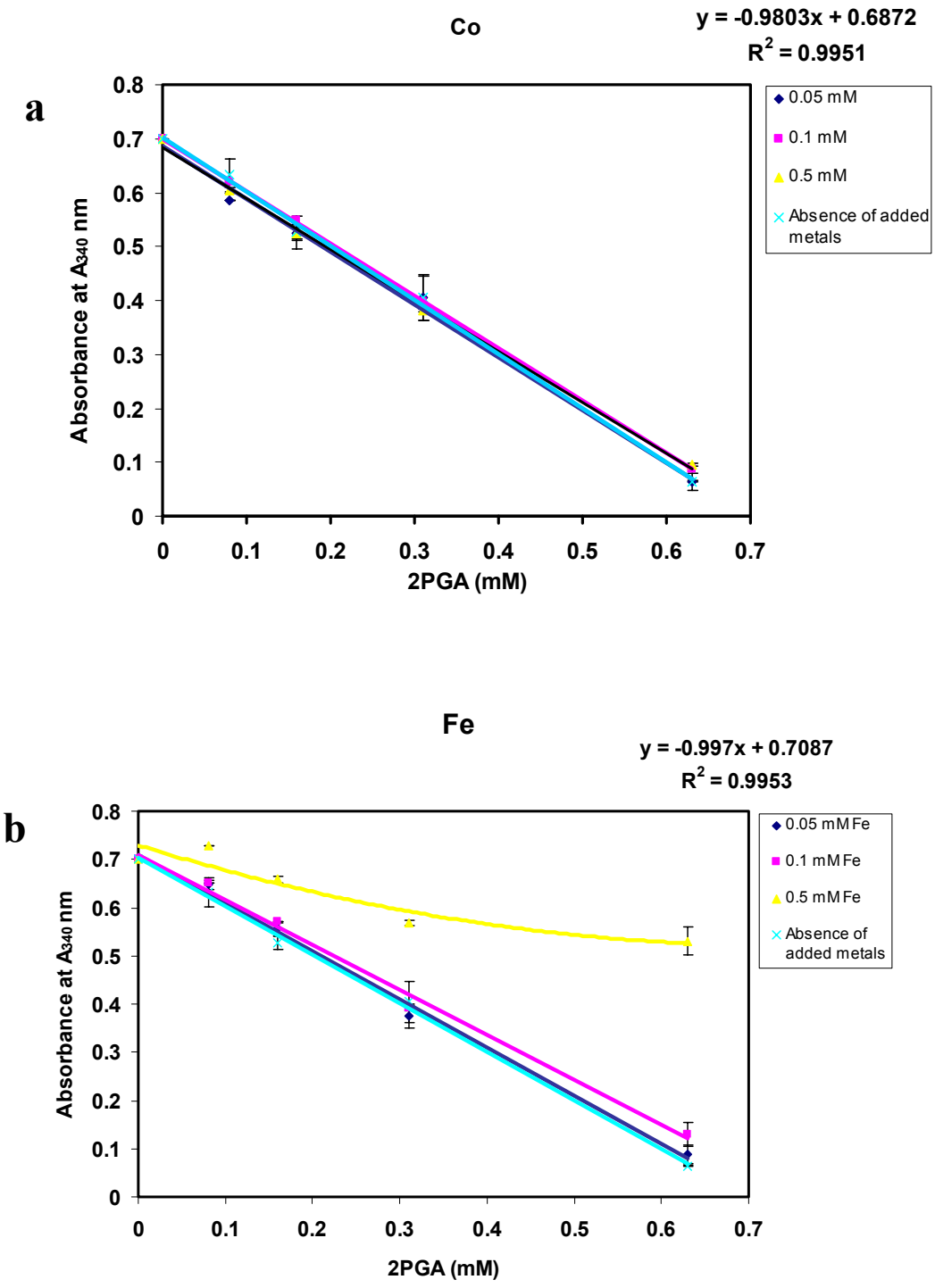


Figure 6.10 Calibration standards in the presence of different concentrations of a) Co^{2+} and b) Fe^{2+} for the coupling enzymes. The rest of the calibration standards are shown in the Appendix I (f). The curve fit and R^2 value were only shown for the 0.05 mM concentration, as this is the calibration standard used for 2PGA concentration measurement.

b) Assay optimisation: Enzyme concentrations, substrate concentrations and time of incubation

Development of the new coupled discontinuous enzyme assay required a number of preliminary experiments in which three important parameters had to be validated: enzyme concentrations, substrate concentrations and length of incubation for the reaction of the enzyme and the substrate to occur. It is often difficult to standardise all the parameters at the same time, hence every single one of them required a 'try and error' test to select the best condition for the assay. Enzyme concentration, in any type of assay, is crucial to be determined. Often adding too much enzyme would lead to very swift attainment of equilibrium, even before the cuvette or plate is placed in the spectrophotometer, resulting in samples all having the same concentration of 2PGA. On the contrary, adding too little enzyme would result in concentrations of 2PGA that would be too small to be measured accurately (Price and Nairn 2009). Although the discontinuous assay is not an absolute quantitation, and is relative to the standard curves generated, it is necessary to ensure that the concentration of the enzyme is optimal to provide meaningful results from the analysis. Reliable measurements also depend on the amount of substrate and the time of incubation for the reaction to occur. In this section, some preliminary experiments which were carried out in the presence of a range of enzyme and substrate concentrations, and at variable time of incubation will be discussed. It is necessary to note that this analysis was carried out using iPGAM from P1 from the ion-exchange column, in the absence of EDTA or any metal chelating agents.

Initially, the reaction was carried out in the presence of 0.1 mg/ml, 0.2 mg/ml, 0.5 mg/ml and 1 mg/ml of iPGAM in the presence of 1 mM Co^{2+}, with 10 mM 3PGA as the original concentration in the first tube (Johnson and Price 1988). The starting absorbance in the plate reader well was standardised to be ~0.7 by adjusting the NADH concentration. As shown in Figure 6.11, the absorbance values of all the samples were very similar to 0.7, thus indicating that the concentration of 2PGA in tube 1 was insufficient. It was learnt later (see section 6.5a) that the concentration of 2PGA from an initial concentration of 10 mM 3PGA was too low to measure reproducibly. The 2PGA concentration at equilibrium was only 7.56% from

the original 3PGA concentration, which means only 0.756 mM 2PGA was present in the original mixture in tube 1. Taking into account the 20-fold dilution made after quenching, only 0.038 mM 2PGA was present in the well during the absorbance measurement. Although this value is still in the range that could be determined from the calibration standards (see Figure 6.10), it is too low to show any significant changes at the different enzyme concentrations.

The binding of a substrate to an enzyme, followed by a catalytic step correspond to two major stages in an enzyme reaction (Figure 6.12). The principle of the reaction lies in the concentrations of substrate and product, which would only change as the reaction progresses, until equilibrium is attained. In this initial experiment, the amount of 2PGA formed was obtained in the presence of 1 mM Co^{2+}, which is known to maximally activate iPGAM in the continuous assay (see section 6.3b) and later found to speed up the time to reach the equilibrium (see section 6.5b). If the time of incubation is too long, all samples, regardless of the metal present, would exhibit similar 2PGA concentrations, as the reaction attained equilibrium. In this respect, besides the enzyme concentration, it is vital to determine the time required for the reaction to be measured before the equilibrium is attained, so that the concentration of 2PGA in all samples could be distinguished. In addition, a sufficient initial 3PGA concentration has to be determined, to ensure an adequate signal that could be monitored using the calibration standards. It is noteworthy that the control samples in the absence of iPGAM or substrate showed no oxidation of NADH.

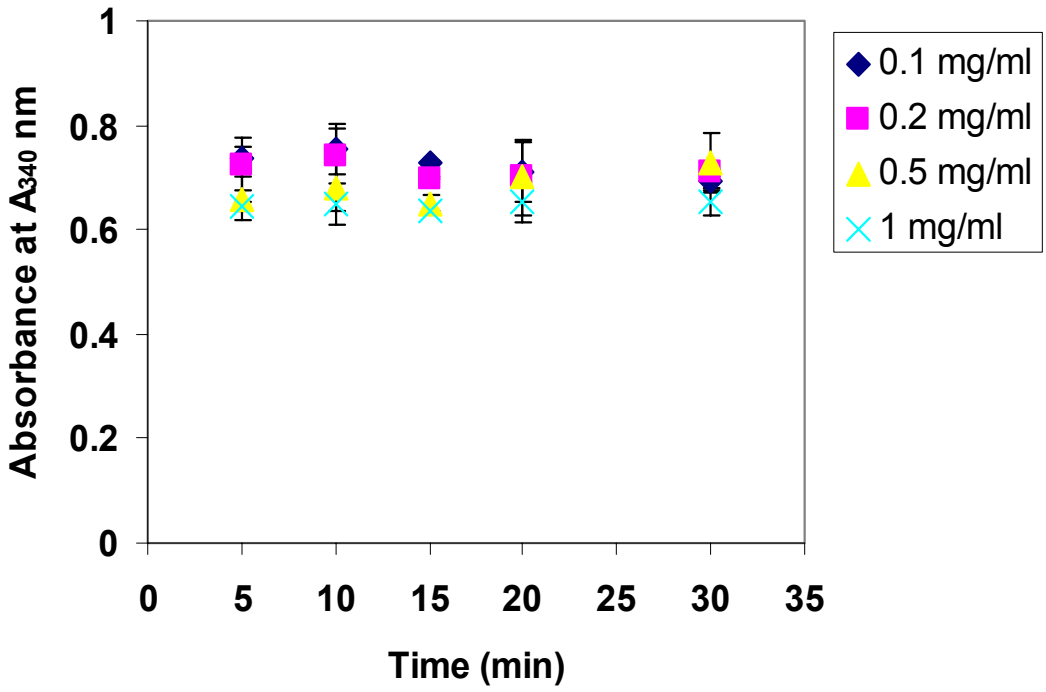


Figure 6.11 Effect of *Lm*iPGAM concentrations on iPGAM activity in the discontinuous assay. Absorbance reading at A_{340} nm exhibited insufficient concentration of 2PGA in the reaction mixture, indicating the requirement for higher initial concentrations of 3PGA.

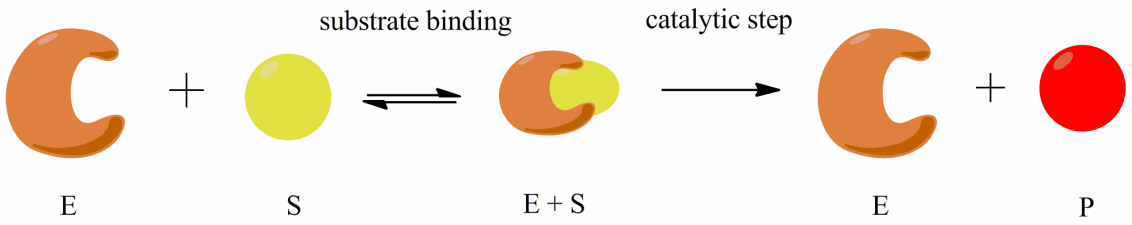


Figure 6.12 The mechanism of an enzyme catalysed reaction. The substrate (S) binds into an enzyme (E) and subsequently produces a product (P).

Subsequently, the effect of 3PGA concentration was studied in an attempt to observe changes in the absorbance. The experiment was conducted in a similar manner, albeit with different initial concentrations of 3PGA; 30 mM, 50 mM, 70 mM and 100 mM with a fixed enzyme concentration of 0.2 mg/ml. The protocol for the activity measurement was as described in chapter 2. As depicted in Figure 6.13, increased initial concentrations of 3PGA resulted in decreased absorbance of NADH. Since 50 mM 3PGA exhibited sufficient signal for the analysis, the next experiments were conducted with 50 mM 3PGA as the initial concentration. The method of quenching was also validated by heating iPGAM at 100°C for 2 min prior to equilibration with 3PGA, whereby no decrease in absorbance was observed, indicating the suitability of the method.

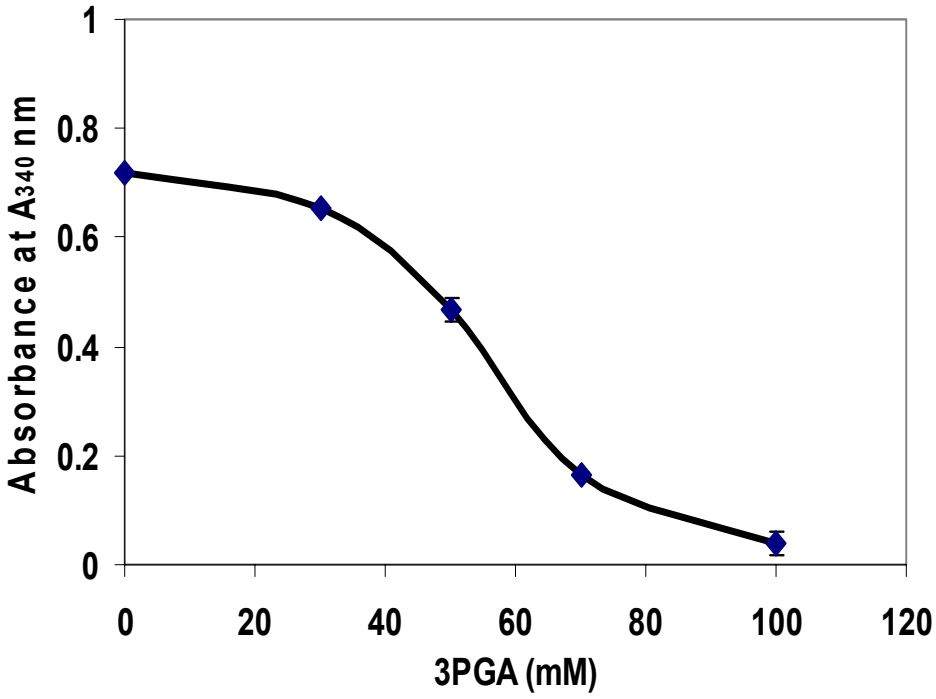


Figure 6.13 Effect of 3PGA concentrations on *Lm*iPGAM activity in the discontinuous assay. Absorbance reading at A_{340} nm showing significant effect of different 3PGA concentrations on the assay system. Errors were calculated but some of the values are too small to show up.

With the purpose of determining the optimum incubation time before the reaction attained equilibrium, a control experiment was set up with iPGAM in the absence and presence of 1 mM Co^{2+}, and the reaction was started with the addition of 50 mM 3PGA. An aliquot of sample was taken at 5, 10, 15, 20, 30, 60 and 120 min after the start of the reaction and quenched with rapid heating for 2 min at 100°C. The differences of the absorbance reading between the two samples are depicted in Figure 6.14. The striking observations are that the decrease in absorbance of NADH is more rapid in the presence of 1 mM Co^{2+}, and levelled off after 30 min incubation. This reflects the ability of Co^{2+}, as determined in the continuous coupled assay system, to hyperactivate the enzyme. While it was essential in this experiment to stop the reaction at different time points to observe the rate of production of 2PGA which is clearly different in the two samples, the concentration of the product would be similar once equilibrium had been reached. An equilibrium level, indicating similar 2PGA concentrations formed, was attained in both samples after approximately 30 min incubation, irrespective of the amount of enzyme activity. In a different albeit similar conditions, Johnson and Price 1988 reported that the final concentrations of 2PGA formed from wheat germ and *Aspergillus nidulans* iPGAMs, and surprisingly rabbit muscle dPGAM were also similar in all cases, after more than 20 min of incubation. Therefore, it is crucial to "catch" the reaction after an appropriate incubation time, to ensure that the enzyme activity is distinguishable between the different samples before the reaction completes. The hyperbolic rate of the reaction within the time interval before the equilibrium phase indicated the suitability of this assay. Consequently, a 15 min incubation time was selected as the time of incubation for the next experiment.

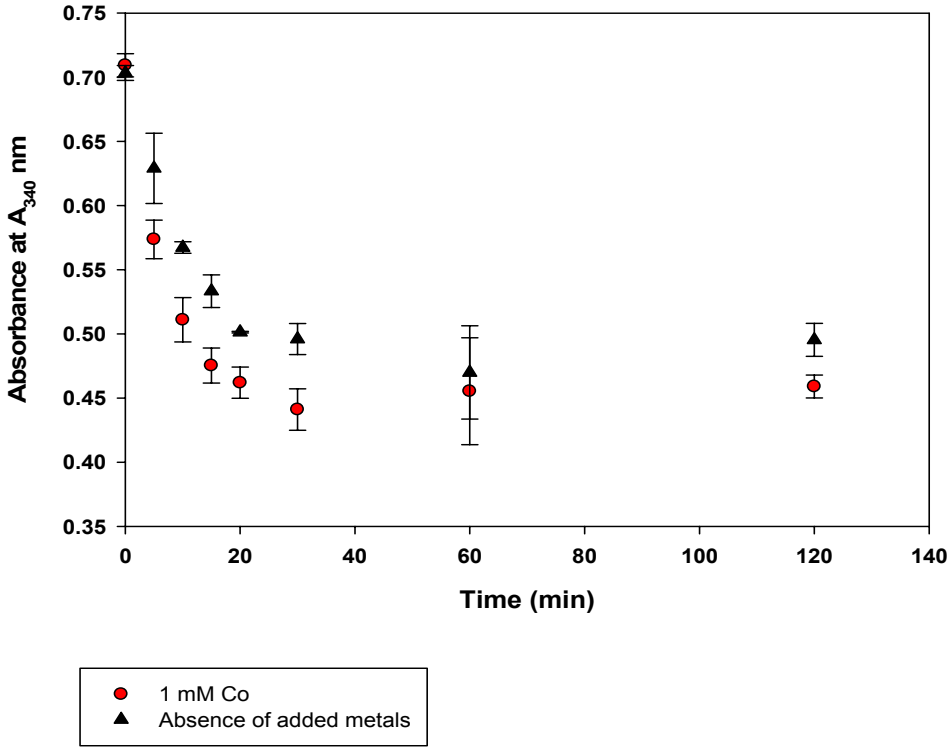


Figure 6.14 Effect of incubation time on *Lm*iPGAM activity in the discontinuous assay. Absorbance reading at A_{340} nm showing significant effect of different times of incubation for iPGAM with the substrate 3PGA. The experiment was conducted in the presence and absence of 1 mM Co^{2+}.

6.5 Enzymatic activities in the presence of different metal ions

The requirement for divalent metal ions by iPGAMs from a range of different organisms has been elucidated previously (Watabe and Freese 1979; Chevalier et al. 2000; Guerra et al. 2004; Raverdy et al. 2007). The metal ions required for each of these iPGAMs however, varies from one organism to another as discussed in the next section. Trypanosomatid iPGAMs have been shown to possess a great affinity for Co^{2+}, as elucidated in chapter 5. Nevertheless, it is plausible that iPGAM, in the presence of Co^{2+}, experiences hyperactivation and the phenomenon observed *in vitro* is not a true reflection of the condition *in vivo*. The optimisation of the activity assay in a discontinuous fashion made it possible to assay iPGAM activity with minimal metal interference on the coupling enzymes. The next section will focus on the activities of iPGAM in the presence of a range of divalent metal ions, as measured by

the concentration of 2PGA. This may be analysed in two ways: (i) the effects of metals at different concentrations on iPGAM activity after 15 min reaction with 3PGA, and (ii) the time required to reach half of the 2PGA concentration at equilibrium in the presence of different metals.

a) Effects of metals at different concentrations towards *Lm*iPGAM activity

As a member of the metal-dependent alkaline phosphatase superfamily, iPGAM evolved as an enzyme that requires metals at the active site. Other members of the family that have been well characterised such as *E. coli* alkaline phosphatase (Kim and Wyckoff 1991), *Bacillus* iPGAM (Kuhn et al. 1995; Nukui et al. 2007) and nematode iPGAM (Raverdy et al. 2007) require Zn^{2+}, Mn^{2+} and Mg^{2+}, respectively. *Lm*iPGAM and other trypanosomatid iPGAMs however appear to function with a one-metal mechanism (Nowicki et al. 2009) in which Co^{2+} performs well as the proposed required metal. As shown in Figure 1.6 (chapter 1), *Lm*iPGAM possesses two metal sites, with metal site 1 (M1) having a greater affinity for Co^{2+}, while M2 by contrast, exhibited poor affinity for the metal. Moreover, Ser75 that is coordinated with the M2 site and involved in catalysis adopts an unfavourable position for phospho transfer if site M2 is occupied by metal. Based on these observations, it is plausible that M1 site occupation is essential for iPGAM activity, while M2 might play a role as a metal regulatory site for *Lm*iPGAM. Nevertheless, with the low affinity for Co^{2+} (or the biologically relevant metal) of the M2 site, the inhibition could only be observed at relatively high metal concentrations. As elucidated in chapter 5, the highest concentration measured in the cytosol of the parasite is Mg^{2+} with 0.2 mM, hence the effects of metals should be tested in conditions close to those resembling the *in vivo* surroundings.

In this section, iPGAM activity was monitored in the presence of biologically relevant concentrations of divalent metals, ranging from 0.01 mM to 1 mM. Altogether, eight metals were tested; Ca^{2+}, Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}, following the protocol as described in chapter 2. The incubation time for iPGAM with the metals is 1 h, prior to the reaction between the enzyme and substrate 3PGA, which was executed for 15 min. It is noteworthy that the metals

were not removed from the iPGAM samples, hence some residual effects of the metals could still be observed from the control experiment. These conditions would provide better mimicking of the *in vivo* surroundings, where divalent metals would also be present at biologically relevant concentrations.

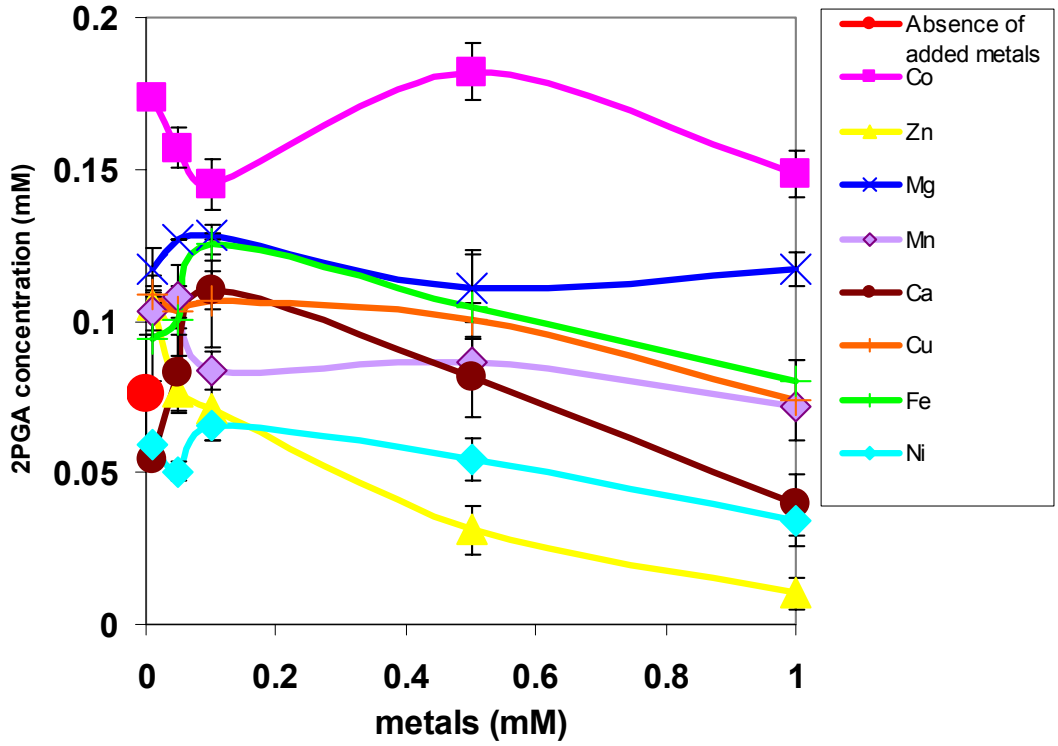


Figure 6.15 Effects of divalent metal ions at different concentrations on *Lm*iPGAM activity. The concentrations of 2PGA formed are those in the plate reader well before correction of the 20-fold dilution.

As mentioned previously, the activity of iPGAM was monitored by the concentration of 2PGA in the original reaction in tube 1. Figure 6.15 depicts the concentration of 2PGA in each sample containing individual metals at different concentrations. All metals at low concentrations, except for Ca^{2+} and Ni^{2+}, stimulated the concentration of 2PGA compared to the native sample in the absence of added metals (red dot), which acted as the control experiment. The trends however, varied greatly among the different metals. Co^{2+} exhibited the strongest activation effects on iPGAM, increasing the activity until 0.5 mM, before decreasing the activity when it reached 1 mM. This observation, in line with the previous hypothesis, suggested that

Co^{2+} activated iPGAM to a certain concentration, approximately 0.5 mM, where primarily metal site 1 is occupied by Co^{2+}. Since M2, as suggested from Nowicki et al. 2009, is the regulatory site albeit a weaker binding site for Co^{2+}, the inhibitory effects would not have been seen until 1 mM concentration, as observed in Figure 6.15. It is anticipated that the activity would decrease as the concentration of Co^{2+} increased. It is noteworthy that the activation of *Lm*iPGAM by Co^{2+} in the discontinuous assay is comparable to the previous continuous assay, although the activation was observed to be less than the former analysis.

A direct observation that could be seen in this experiment is the concentration of 2PGA in the samples, which interestingly is much lower than the initial substrate concentration (50 mM). This is an unexpected result for *Lm*iPGAM, where in the crystal structure, 3PGA and 2PGA are present at equal occupancies (Poonperm 2005; Nowicki et al. 2009). Nevertheless, Johnson and Price 1988 reported that the concentration of 2PGA in the *A. nidulans* iPGAM reaction was only 0.8 mM from the original concentration of 10 mM 3PGA. A similar observation was obtained from the present study, where 3.78 mM 2PGA was measured (after correction of dilution) from the original concentration of 50 mM 3PGA, in the presence of 0.5 mM Co^{2+}. While in the older report, the equilibrium mixture contained 8% 2PGA, the latter experiment resulted in 7.56% 2PGA. It is clear from this result that in solution, the substrate: product ratio at equilibrium is far from 1:1, and in fact is 12:1. The K_m value for 2PGA, which is two-fold lower than 3PGA (Guerra et al. 2004) might explain this observation, as the continuously formed 2PGA with tighter binding to the active site reacted *in situ* with iPGAM to produce more 3PGA.

Among all metals, Mg^{2+} showed a slight activation of iPGAM, across the whole range of concentrations. It is noteworthy that Mg^{2+} is found at the highest concentration of any divalent metal in the protein sample (8 μM), and might contribute to the residual activity found in the sample in the absence of added metals. A similar trend of Mg^{2+} concentration is also observed *in vivo* in the parasite cytosol. It is postulated that Mg^{2+}, by being the most abundant metal in the cell, can maintain the activity of iPGAM at a low level, albeit less than Co^{2+} *in vitro*. This is clearly seen by the regulation of the activity of iPGAM by Mg^{2+} (Figure 6.15), which was

maintained at the same level over a range of Mg^{2+} concentrations up to 1 mM (five times the cytosol concentration of ~200 mM (see chapter 5). Mn^{2+} showed a similar effect, albeit at a lower level of iPGAM activity. Ca^{2+}, Cu^{2+} and Fe^{2+} showed decreased iPGAM activity at higher concentrations. Ni^{2+} and Zn^{2+}, on the other hand, inhibited the enzyme even at low concentrations. Zn^{2+} moreover, showed significant inhibition of the enzyme activity with almost no 2PGA formed when the Zn^{2+} concentration increased to 1 mM.

b) Divalent metals that are activators, inhibitors or have no effect on *Lm*iPGAM

A second experiment was conducted in the presence of various divalent metal ions. This analysis however, was employed in the presence of fixed concentrations of metals at 1 mM. A similar protocol was conducted for all experiments, and a portion of the enzyme catalysed reaction was quenched at 5, 10, 15, 30, 60 and 120 min. It was anticipated that the data would fit a hyperbolic curve, where the exponential phase can estimate the rate of activity in the presence of each metal. This could be performed by determining the time to reach half equilibrium 2PGA concentration, separately for each individual metal. The analysis was similar to the determination of K_m value from the Michaelis-Menten plot.

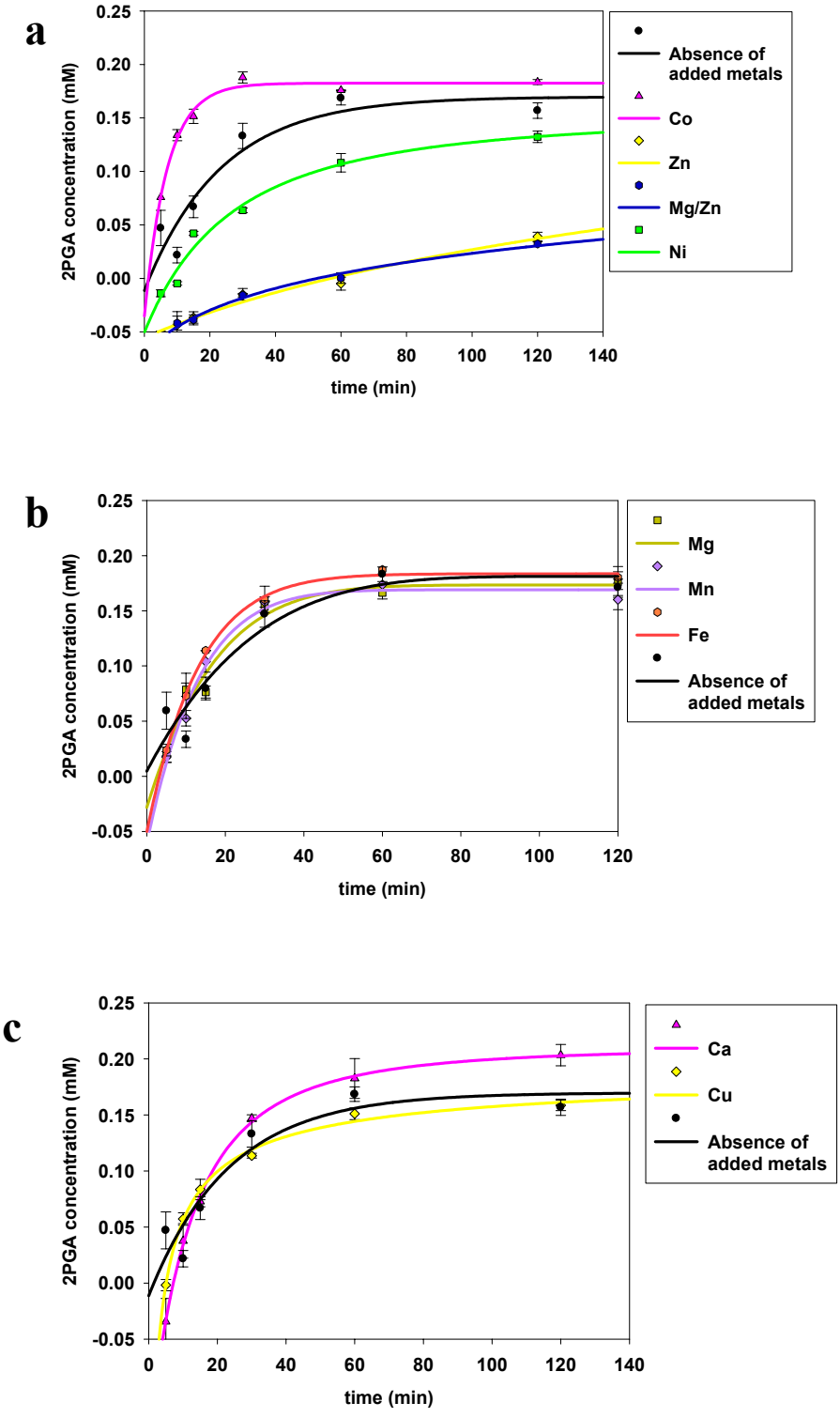


Figure 6.16 Effects of metals towards *Lm*iPGAM activity is shown, and the types of metals are grouped as follows: a) Strongly activating and inhibiting metals b) Slightly activating metals c) Metals that have no effect to *Lm*iPGAM.

Table 6.2 The time required to produce half of the equilibrium 2PGA concentration in the presence of different divalent metal ions at 1 mM.

Divalent metals	Time to reach half [2PGA] at equilibrium (min)
Ca^{2+}	18.5
Co^{2+}	5.66
Cu^{2+}	18.8
Fe^{2+}	12.1
Mg^{2+}	13.8
Mn^{2+}	12.6
Mg^{2+}/Zn^{2+}	Off scale
Ni^{2+}	33.0
Zn^{2+}	Off scale
Absence of added metals	16.2

The results (Figure 6.16) showed that most of the data fit the hyperbolic curves very well, and all the plots were acquired by using the Modified Hyperbola 3 curve fit (Sigma Plot). The original equation; y = a-b/(1+cx)1/d, was derived to obtain the time (x) value to obtain the half maximum 2PGA concentration at equilibrium (y). The results from this experiment categorised the divalent metals into specific groups, based on their effects towards the enzyme activity. Figure 6.16a illustrates the effects of the iPGAM activator Co^{2+}, and the inhibitors Ni^{2+} and Zn^{2+}, compared to the sample in the absence of added metals. The calculated value from the equation showed that samples in the presence of Co^{2+} reached half of the maximum 2PGA concentration at equilibrium almost three-fold faster than in the absence of added metals (see Table 6.2). On the contrary, Ni^{2+} slowed the rate of the reaction by a factor of ~2, and Zn^{2+} slowed the reaction beyond a measurable value. Interestingly, the addition of Mg^{2+} into the samples in the presence of Zn^{2+} did not restore the activity of iPGAM, which remained unambiguously the same as the sample with Zn^{2+} alone. Relating this result to *in vivo* conditions, it is important to note that the total Zn^{2+} concentration in the cytosol is ~6-fold lower than the total

Mg^{2+} concentration, and many of these metal ions might be tightly bound to other macromolecules which are present in the cytosol. Unfortunately, the method used to determine the divalent metal concentrations as explained in chapter 5 does not distinguished between bound or free metals in the analysed sample and hence, the actual concentrations of free metals in the cytosol are not known.

Some of the other divalent metals fell into the slightly activating metals category, specifically Fe^{2+}, Mn^{2+} and Mg^{2+} (Figure 6.16b). These three metals required 12.1, 12.6 and 13.8 min respectively, to reach the half maximum 2PGA equilibrium concentration, compared to the sample in the absence of added metal which required 16.2 min (see Table 6.2). Interestingly, Fe^{2+} and Mg^{2+} (as well as Zn^{2+}) were among the three metals that were found to be abundant in the parasite cytosol (chapter 5). On the other hand, Ca^{2+} and Cu^{2+} do not exhibit significant activating effects on iPGAM activity (Figure 6.16c), and had values close to the control sample, requiring 18.5 min and 18.8 min (Table 6.2) to produce half of the maximum 2PGA equilibrium concentration, respectively.

In summary, it is clear that the selectivity of metals for iPGAM to employ its function is different from one metal, to another. Co^{2+} at all concentrations tested, has greater activating effects on iPGAM, compared to the rest of the metals. Interestingly, Mg^{2+}, being the most abundant metal *in vivo*, could sustain the activity of iPGAM throughout the range of concentrations, albeit to a lower level than Co^{2+}. This result supports the preference of the octahedral geometry of the M1 site for Co^{2+} and Mg^{2+}, and provides the possibility that Co^{2+} *in vitro* could replace Mg^{2+} by occupying the M1 site. Zn^{2+}, on the other hand, is inhibitory toward iPGAM. It may be postulated that Zn^{2+} coordinates the M2 site that possesses tetrahedral geometry, and thereby contributes to the inhibitory effect. A mechanism for the regulation of the activity of iPGAM in the trypanosomatid cytosol is modelled in Figure 6.17. Although the mechanism of the enzyme *in vivo* is still poorly understood, it is plausible that the activation and inhibition of iPGAM are regulated by divalent metals, resulting in the relatively low specific activity as measured in the cytosol of the bloodstream and procyclic forms of the parasite.

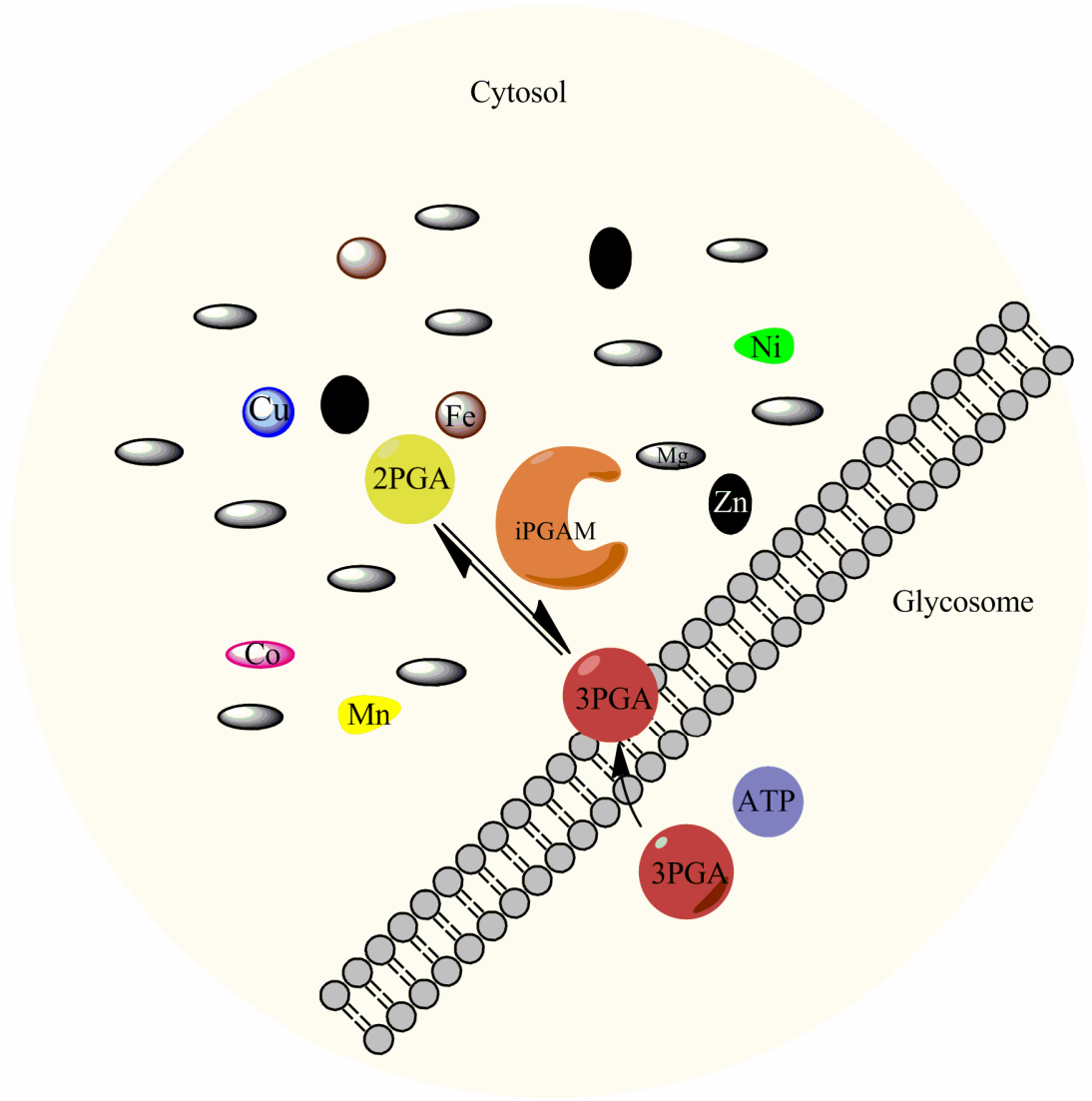


Figure 6.17 The potential for the occupation of the metal sites of trypanosomatid iPGAM in the cytosol of the parasites with Mg^{2+} and Zn^{2+}. Other divalent metals are present at significantly lower concentrations than either Mg^{2+} or Zn^{2+}. However, the concentration of the free metals in the cytosol and glycosome are not known, and the model must remain speculative.

c) Enzymatic activities in the presence of higher concentrations of monovalent cations Na^+ and K^+

Monovalent cations, Na^+ and K^+ have been shown to be essential for regulating the activity or maintaining the structural integrity of enzymes (Page and Di Cera 2006). The crystal structure of *Lm*iPGAM (Nowicki et al. 2009) revealed the presence of a Na^+ ion in the phosphatase domain, although the occupied location is distant from the active site (Figure 1.6 chapter 1). The main role of Na^+ is unclear, but believed to be required for structural integrity. In addition, Na^+ is required to activate a number of enzymes such as fructose 1,6-bisphosphate aldolase (Hall et al. 1999), an enzyme that plays a role in both glycolysis and gluconeogenesis, and tagatose 1,6-bisphosphate aldolase (Hall et al. 2002), which is involved in galactose metabolism. These enzymes however, displayed the pairing of Na^+- Zn^{2+} rather than K^+-Mg^{2+} typical configuration observed in most phosphoryl transfer enzymes (Page and Di Cera 2006). A vital role for K^+ has been identified in a number of phosphoryl transfer enzymes such as pyruvate kinase (Larsen et al. 1998), pyruvate dehydrogenase kinase (Kato et al. 2005), and ribokinase (Andersson and Mowbray 2002; Zhang et al. 2004), which interestingly also have absolute requirements for the divalent metal cation, Mg^{2+} (Page and Di Cera 2006).

Taking into account the requirement for K^+ by pyruvate kinase, this monovalent cation has always been present in excess in the assay buffer, at 50 mM. Na^+ on the other hand, is included throughout the purification procedure and kept in the storage buffer for iPGAM structural integrity, also at 50 mM. Hence, in order to observe potential significant effects of these monovalent cations towards iPGAM activity, the concentrations were increased to be in the range of 100 to 200 mM. The typical condition, which contains 50 mM Na^+ and 50 mM K^+ was included as the control experiment.

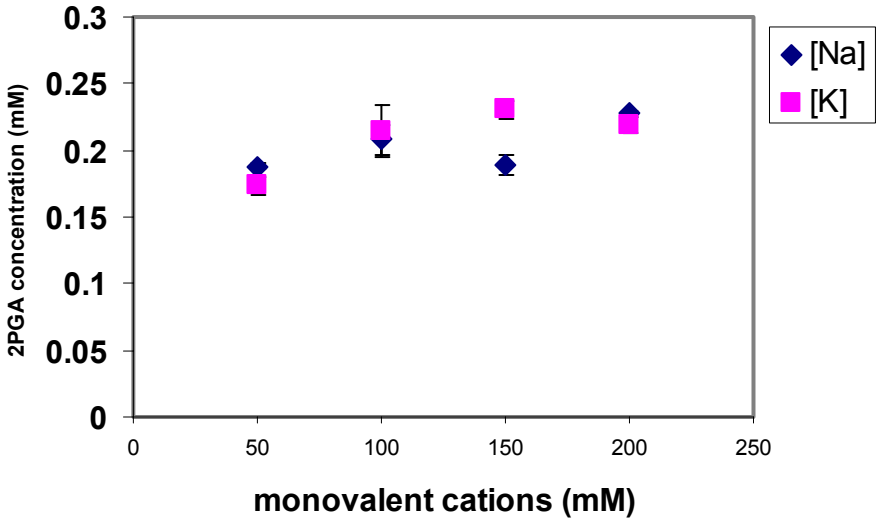


Figure 6.18 The formation of 2PGA in the presence of different concentrations of monovalent cations.

The results indicated that slight activation of the enzyme could be observed with increasing concentrations of K^+, while fluctuation is observed with higher concentrations of Na^+ (Figure 6.18). It is not possible however, at this stage to determine whether these monovalent cation metals are required for iPGAM activity. The slight activation may well be the direct effect of K^+ on pyruvate kinase activity in the second reaction of the discontinuous assay.

6.6 Conclusion

An assay system that could be employed in the presence of metal ions, with minimal metal interference to the other coupling enzymes has been successfully developed. As expected, greater activation was observed in the presence of Co^{2+}, although the activity of iPGAM decreased when the concentration of Co^{2+} exceeded 0.5 mM. The time required for iPGAM in the presence of Co^{2+} to produce half of the maximum 2PGA equilibrium concentration is three-fold faster than in the control experiment (in the absence of added metal), supporting the previous findings that the enzyme could be hyper activated *in vitro*. At equilibrium the concentration of 2PGA was 7.56% of the 3PGA concentration, implying that the reversible mechanism of the enzyme is favouring the reverse reaction. In addressing the "correct" metal *in vivo*,

Fe^{2+}, Mg^{2+} and Mn^{2+} exhibit slight activation of iPGAM, albeit less than Co^{2+}. Mg^{2+} was shown to activate iPGAM across a 100-fold range of concentrations, suggesting that it may continuously support iPGAM activity at a relatively low level *in vivo*. The relative concentrations of these three divalent metal ions in the cytosol were in the order Mg^{2+}>Fe^{2+}>Mn^{2+}. A similar trend of metal concentrations was also observed in the bacterially expressed *Lm*iPGAM.

By contrast, Ni^{2+} and Zn^{2+} exerted inhibitory effects on iPGAM activity. While the concentration of Ni^{2+} in the cytosol is insignificant to satisfy a 1:1 stoichiometry of metal: iPGAM, Zn^{2+} appeared to be present in excess concentration to fulfil the stoichiometry. Interestingly, the M1 site in the active site of iPGAM has octahedral geometry, indicating its preferences for Co^{2+} or Mg^{2+}, while the M2 site possesses tetrahedral geometry, with a tendency to be occupied by Zn^{2+}. Taken together, it is plausible that Mg^{2+} and Zn^{2+} are the native metals which regulate iPGAM activity *in vivo,* albeit keeping it functioning at a relatively low level. However, it is noteworthy that if all Zn^{2+} which is present in the cytosol is free, the activity of iPGAM would be inhibited, which is believed to be proven otherwise. Some of Zn^{2+} might be bound tightly to other macromolecules, which means that it, with ~6-fold lower concentration than Mg^{2+}, may play a role in regulating the activity of iPGAM. The mechanism and behaviour of iPGAM in its *in vivo* surroundings is yet to be explored in detail, but it may well turn out to have the potential to be a key modulator in controlling the glycolytic flux in trypanosomes (Achcar et al. 2012).

Chapter 7:

Preliminary Inhibitor Screens for *Lm*iPGAM

7.1 Aims

The search for drug candidates in order to inhibit target proteins has been one of the focal points in many trypanosomatid parasite studies (Albert et al. 2005; Wiese et al. 2009; Cavalli et al. 2010; Gupta et al. 2011). This is particularly highlighted for glycolytic enzymes, as a number of them have been validated as promising drug targets, including iPGAM (Albert et al. 2005). At present, a robust high-throughput screening programme of trypanosomatid glycolytic enzymes including *Lm*iPGAM is being conducted in the NIH, Bethesda, USA to identify potent inhibitors. While high-throughput screens provide an exceptional method for identifying drug candidates (Taylor et al. 2008; Agresti et al. 2010), a complementary *in silico* approach with similar objectives can also be employed by using virtual screening programs. Both analyses, which currently are being employed simultaneously by the NIH and Edinburgh research groups, are aimed at elucidating the ability of specific compounds to inhibit enzyme activity *in vitro*. It will be in subsequent stages, where the potency of inhibitors on cultured parasites and on parasite/animal models will be determined.

While the *Lm*iPGAM closed-form structure is available from structural studies, the poorly accessible catalytic site (see chapter 4) made it difficult to perform structure-based drug design experiments. Since the open-form *T. brucei* structure has only recently become available (Mercaldi et al. 2012), the virtual screening analysis had to be carried out with either: i) a model of the open-form based on *Ba*iPGAM open conformation, or ii) based on the known ligands. Attempts to use the model were unsuccessful because the open-form model was too large, and this may well be the consequence of the unfavourable conformation of the open-form *Ba*iPGAM (Mercaldi et al. 2012). Hence, the second approach is a better choice. This approach was employed by initialising similar compounds that resemble the original ligands (substrate 3PGA and/or product 2PGA) which occupy the catalytic site before performing any docking and scoring experiments. In this chapter, some preliminary

results based on this screening approach, as well as high-throughput screening and inhibitory analysis will be discussed.

7.2 Introduction

The interactions between proteins and ligands have always been the basis of any drug design implementation. In this case, the fundamental principle involving explicit biochemical interactions between the amino acid side chains and the binding molecules were manipulated for the purpose of altering the activity of the protein. Enzymes in particular, in a normal biochemical reaction, interact with the substrate which acts as a reactant. These interactions however, may be disrupted by introducing another molecule which possesses the ability to inhibit the binding of the substrate to the enzyme, either by competing or altering the enzyme's mechanism (Klebe 2009). iPGAM in particular, is folded into two domains and competitive inhibitions may occur in a reversible or irreversible fashion, in which non-covalent (hydrogen, ionic or hydrophobic) or covalent (specific interactions between reactive functional groups with amino acid side chains) interactions between the molecule and the residues in the active are involved. For reversible inhibition, the inhibitor may also be competing with the substrate to bind to the active site. Another possibility of inhibition is the presence of any molecules that may not be present in the active site, but lie close enough to the hinge regions and thus affect the movement of the two domains. This is known as non-competitive inhibition. The possible inhibition mechanism for iPGAM is pictured in Figure 7.1. Often too, an activator, which increases the activity of the enzyme, is also taken into consideration in designing drugs for altering metabolic reactions (Temperini et al. 2008).

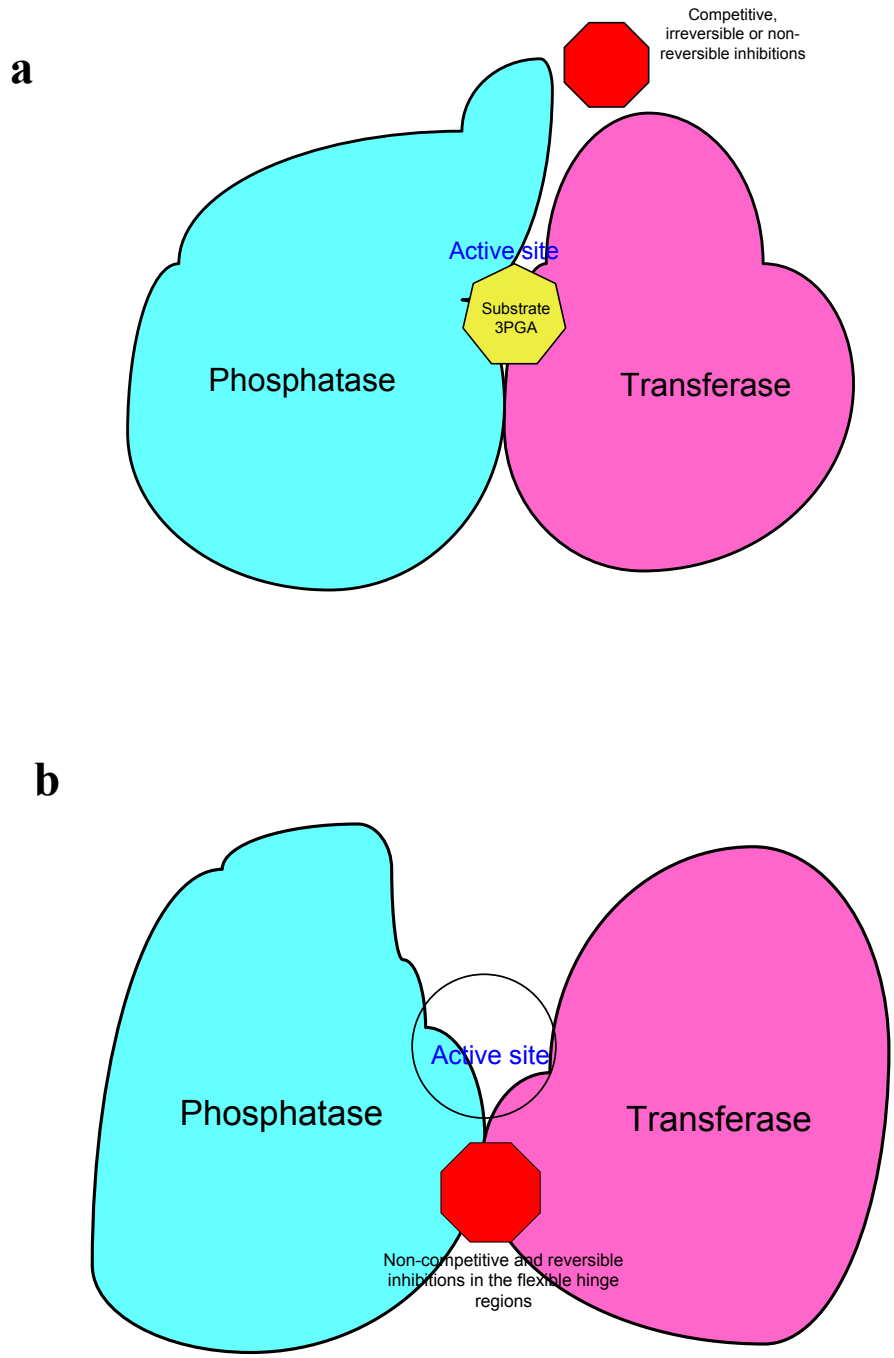


Figure 7.1 The possibilities for iPGAM inhibition. a) Competitive inhibition which may occur either in a reversible or irreversible fashion in the active site and b) Non-competitive inhibition which may occur in a reversible fashion, notably in the hinge regions that control the movement of the two flexible domains.

Over the years, the majority of drug design strategies have involved two possible routes: high-throughput or virtual screens. The first possibility allows experimentally-determined analyses to be conducted through biochemical and pharmacological laboratory tests. The outcome would be the basis for subsequent analyses to find specific interactions between the proteins and the compounds, through another round of biophysical and structural studies. The latter alternative, on the other hand, involves fast and rapid searches of chemical structures in large compound libraries, where the most likely molecules are evaluated to determine their possibility of binding to the target protein (Rester 2008; Rollinger et al. 2008). This can be related to the majority of drugs showing pharmacological effects by explicit interactions with the target proteins, which is often understood based on a variety of models: lock-and-key (Fischer 1894), induced-fit (Koshland 1958), and conformational selection or population shift (Kumar et al. 2000; Boehr et al. 2009). By analysing the chemical structure and dynamic conformational properties of these molecules, the effects on the receptor protein could be predicted (Shim and Mackerell 2011). Subsequently, a series of laboratory experiments would be employed to obtain the experimental outcome for the predicted interactions. Both the high-throughput and virtual screening approaches are shown in Figure 7.2. It is noteworthy that at present, iPGAM inhibitors are still being sought, with no well-established inhibitors except for the non-specific metal chelators reported for the enzyme. As mentioned above, both strategies are being used to search for potent inhibitors for iPGAM, and will be discussed in the later sections.

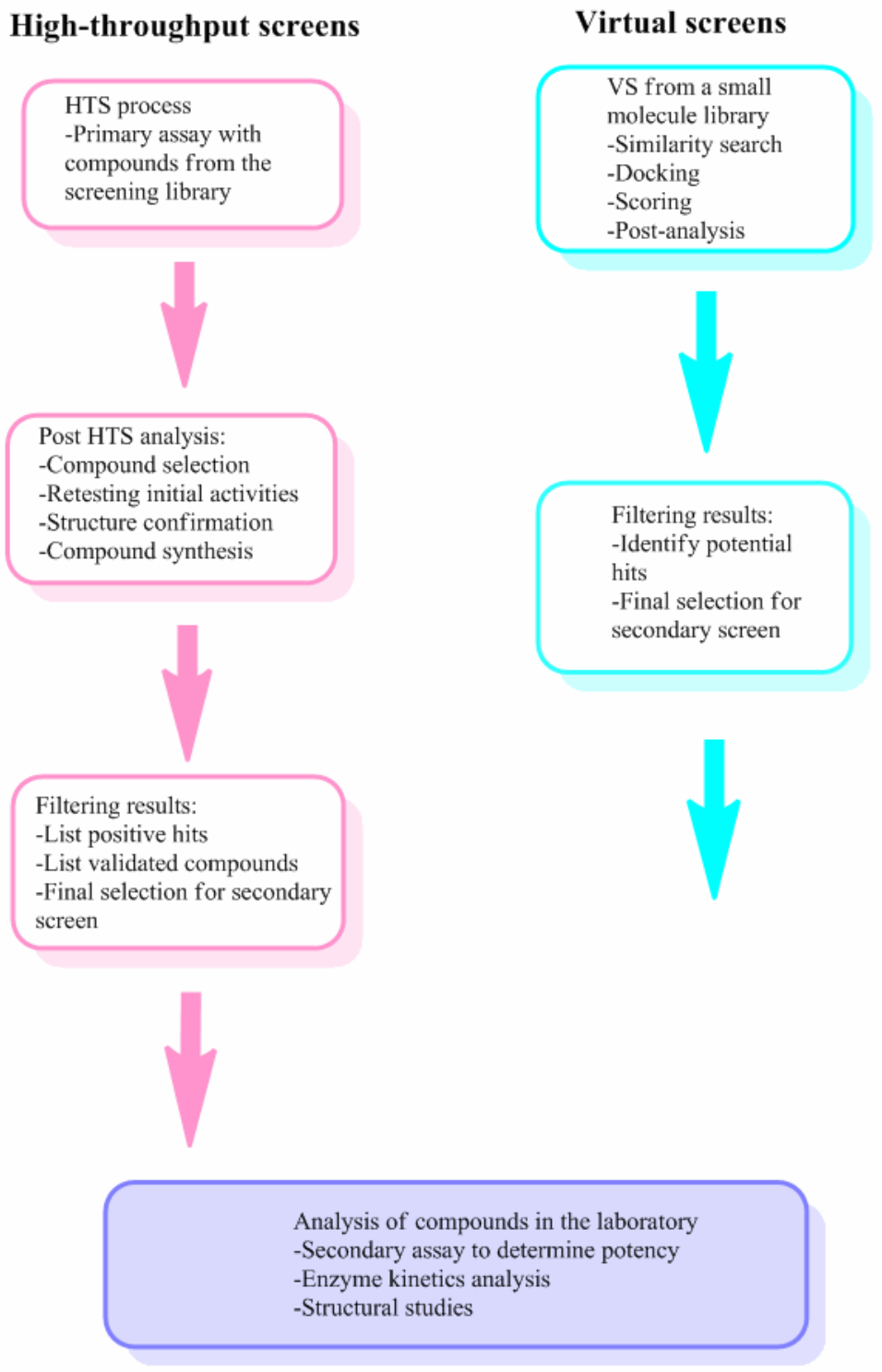


Figure 7.2 The steps involved in high-throughput and virtual screening approaches.

7.3 *In silico* virtual screening strategy

Since virtual screens provide fast, convenient and economical strategies for selecting possible drug candidates, this approach was considered for the initial *Lm*iPGAM drug discovery analysis. However, as discussed in the earlier section, the major obstacle of this approach for *Lm*iPGAM is its poorly-accessible catalytic site. For this reason, the experimental approach has to be designed based on the chemical and structural characteristics of the substrate or product. This is the common route in virtual screening analysis, in cases where the structure of the target molecule is unavailable or suffers poor accessibility to the active site. As this is the case for *Lm*iPGAM, the screening was initially employed following the ligand-based method, before a subsequent round of docking and scoring experiments was conducted with the structure-based method. Both the ligand-based and structure-based drug design approaches are shown in Figure 7.3.

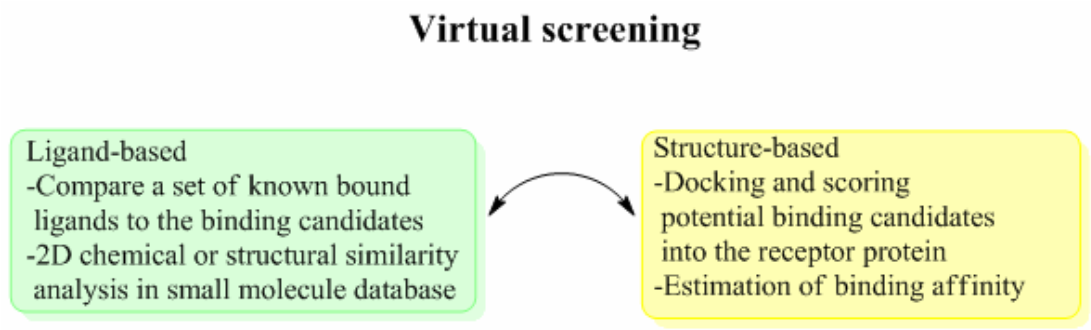


Figure 7.3 Different approaches in a virtual screening process.

a) Ligand-based screens: UFSRAT for identifying small molecule mimics of the substrate and/or product

The ligand-based virtual screen method hinges on the similarities of the query molecule to the compounds that are stored in any compound library or database. The similarities are often regarded as a preamble to the capability of the compounds to mimic the structure and activity of the query molecule. This is particularly interesting for *Lmi*PGAM, as the small and poorly accessible active site results in difficulties for predicting the interactions with any drug-like compounds. It is hence important to find molecules which are similar to the substrate 3PGA or product 2PGA to be inserted into the active site prior to conducting any docking and scoring experiments. For the purpose of identifying the potential candidates, UFSRAT (Ultra Fast Shape Recognition with Atom Types) (Shave 2010) algorithm was employed, utilising both 3PGA and 2PGA as the query molecules (Figure 7.4).

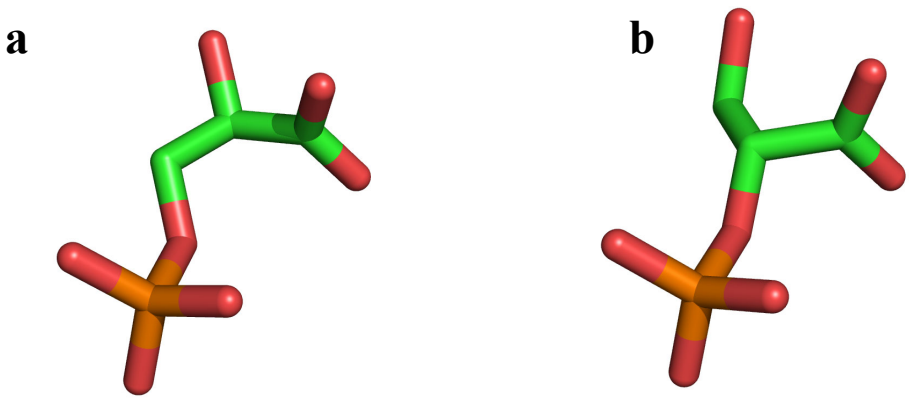


Figure 7.4 The query molecules which were used for UFSRAT: a) 3PGA and b) 2PGA molecules. The red atom refers to oxygen (O), the green refers to carbon (C) and the orange refers to the phosphate atom (P).

The UFSRAT algorithm is based not only on overall shape similarities, but also considers the types and molecular topology of the existing atoms. In this case, the geometrical distribution of atoms was divided into four categories in which consider: all atom types, hydrophobic, hydrogen acceptor or hydrogen donor atoms (Adie 2010; Shave 2010). Taken these descriptors into consideration, a search was employed through a large multi conformer library comprising 4,853,000 molecules from various suppliers: Chembridge, Asinex, Maybridge, Enamine, LifeChemicals (InfLab) Specs, InterBioScreen and ChemDiv, to identify candidates similar to the query molecule. Evaluation was performed based on the scoring function, where a molecule with a score of 0 is the least similar to the query molecule, and a score near to 1 suggests a closer resemblance. The top five most similar candidates for 2PGA and 3PGA are shown in Figure 7.5a and b, respectively, and the complete results are listed in the Appendix II sections. For the subsequent docking experiment, the most favourable compounds which were similar to 2PGA were utilised. This was performed due to the experimentally determined K_m value for 2PGA (0.11 ± 0.03 mM) which was lower than 3PGA (0.27 ± 0.02 mM) (Guerra et al. 2004), suggesting tighter binding to *Lm*iPGAM.

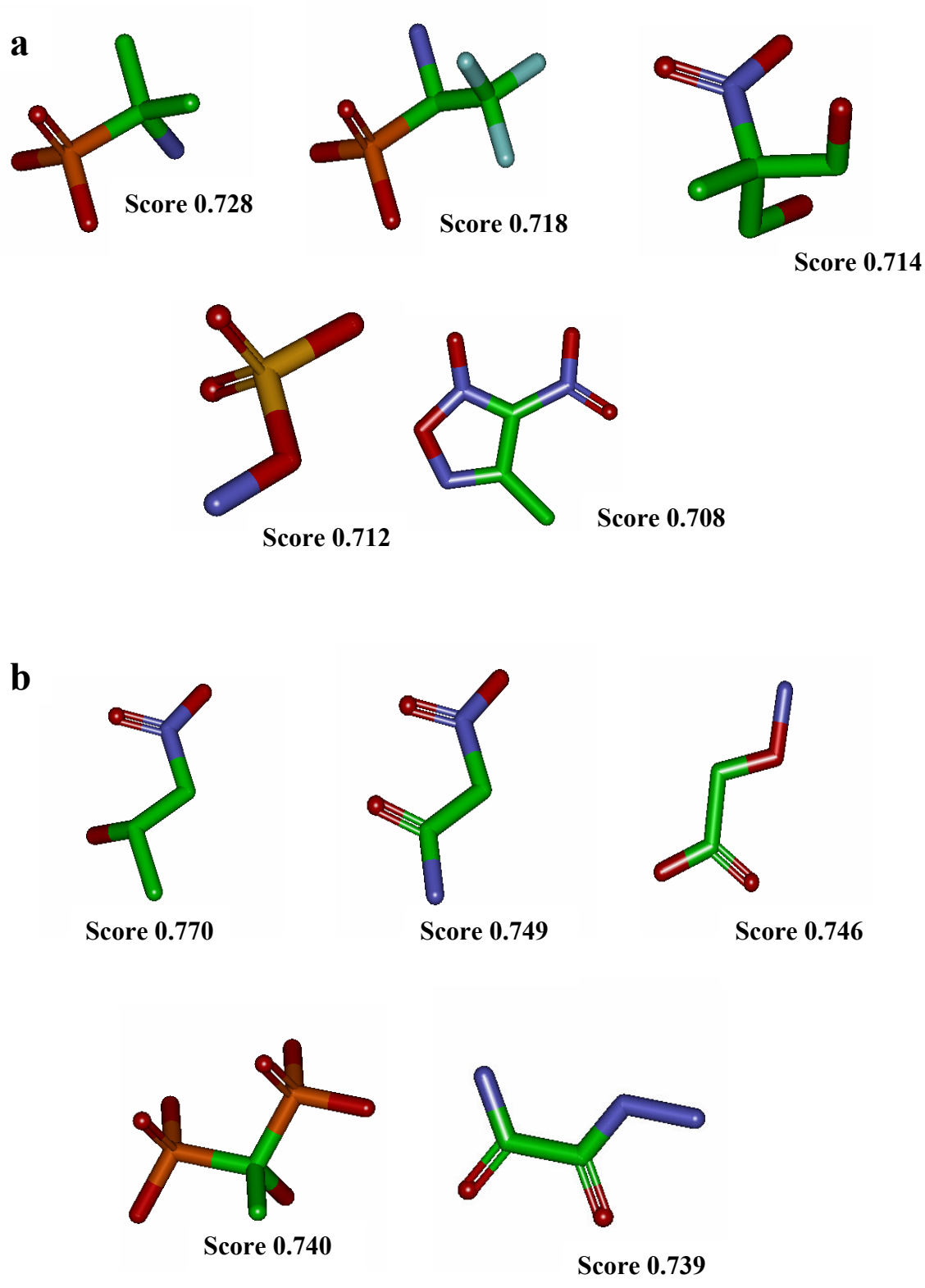


Figure 7.5 The top five highest-scored compounds which were similar to a) 2PGA and b) 3PGA. The atom colourings are identical to Figure 7.4, with the addition of the cyan colouring for fluorine (F) and the blue colouring for nitrogen (N).

b) Structure-based screens: Docking and scoring algorithms

Docking and scoring functions are widely applied in virtual screening for three important purposes: i) to envision the binding mode of a known ligand for the target molecule, ii) to pinpoint new binding candidates for protein-ligand interactions and, iii) to predict the binding affinities of the candidates (Leach et al. 2006). Often a particular interest for enzymatic analysis is to observe possible interactions between the receptor protein and small molecules which could possibly modify the catalytic functions. Nowadays, most programs enable ligand flexibility during docking experiments, with rigid or almost rigid protein structure (Leach et al. 2006), before the docking results were evaluated through a scoring function. A few docking programs such as GLIDE, Gold, and ICM (Perola et al. 2004), and scoring algorithms (GlideScore and ChemScore), as well as virtual screening with combination of both, such as AutoDock4 (Morris et al. 1998; Huey et al. 2007; Morris et al. 2009), AutoDock Vina (Trott and Olson 2010), and LIDAEUS (Taylor et al. 2008) are some of the programs available for predicting protein-ligand interactions. In this study, virtual screening programs combining both algorithms were used: AutoDock4 and AutoDock Vina using the outcome from UFSRAT, as well as a combination of a few virtual screens programs performed iteratively: AutoDock 4, AutoDock Vina, and LIDAEUS (LIgand Discovery At Edinburgh UniverSity) based on the in-house relational small-molecule database EDULISS (EDinburgh University LIgand Selection System) (Hsin et al. 2011).

i) AutoDock4 and AutoDock Vina

AutoDock4 is a computational docking program which is performed on the basis of two important methods: i) exploration of the conformational spaces available in the targeted region and ii) empirical and semi-empirical force fields to evaluate the energetics of the conformations (Huey et al. 2007). Two main approaches are included in the program: i) AutoDock employs ligand docking to a set of grids representing the target receptor and ii) Autogrid which pre-calculates the grids. Meanwhile, AutoDock Vina is the updated version of the program, which surpasses the previous functions of AutoDock4 in atom type selection and grid-map

calculations. In the Vina version, the grid maps are automatically calculated, thereby showing the results transparently to the users (Trott and Olson 2010). This further improves the accuracy and speed of the binding-mode prediction.

Both AutoDock and Vina analyses were carried out for *Lm*iPGAM, utilising the closed-form structure as the target receptor, against the UFSRAT results for 2PGA (Figure 7.6). The hits were listed based on individual scores, and the top five compounds are shown in Table 7.1a and b. The full list is shown in the Appendix II section. The majority of results possess the criteria that are commonly accepted by Lipinski's rule of 5 for molecules' drug-likeliness (Lipinski et al. 1997), where: i) the molecular mass ≤ 500 Da, ii) ≤ 5 hydrogen bond donors, iii) ≤ 10 hydrogen bond acceptors and iv) MLogP (the octanol-water partition coefficient that measures the hydrophobicity of compounds) that is ≤ 4.15 (Moriguchi et al. 1992) (Taylor et al. 2008). Both the AutoDock and Vina results were individually assessed by the scores, where the lower the values, the better the scores. A small number of compounds was purchased (see section 7.5) for the subsequent inhibition analysis.

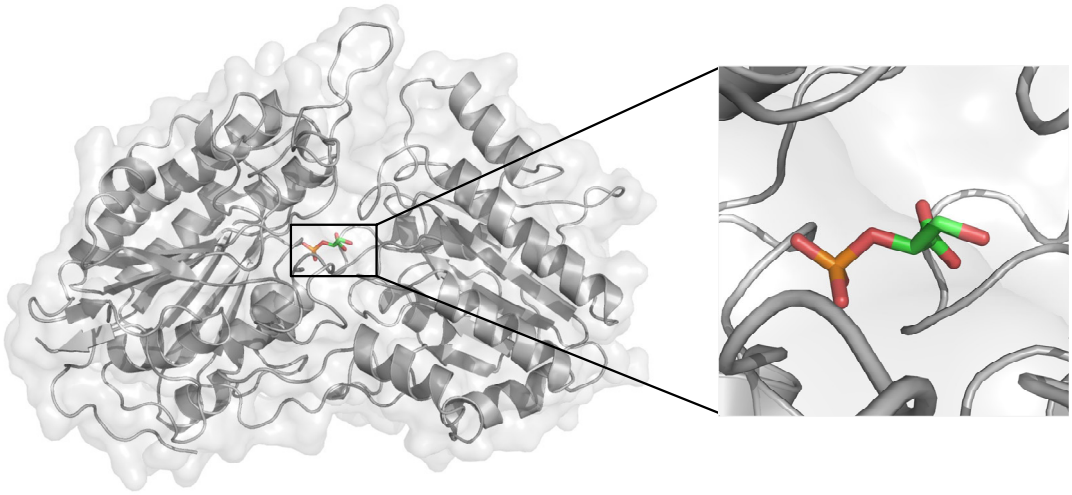


Figure 7.6 The closed-form crystal structure of *Lm*iPGAM with 2PGA in the active site.

Table 7.1 The top five hit molecules from a) AutoDock and b) AutoDock Vina

a

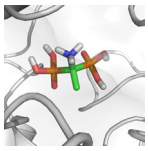
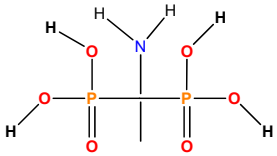
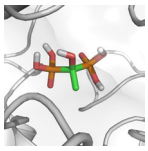
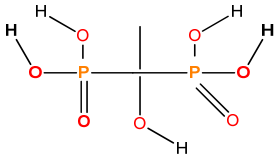
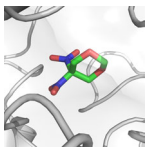
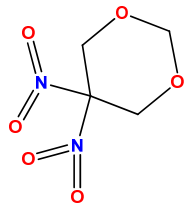
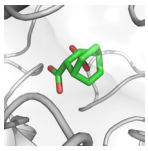
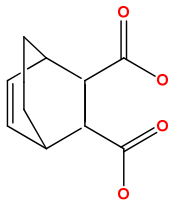
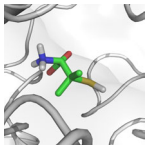
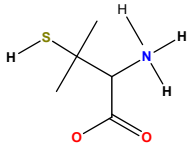
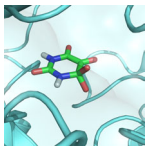
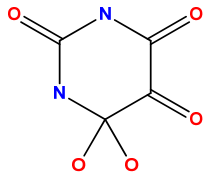
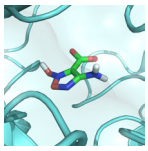
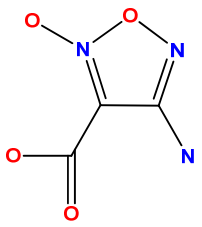
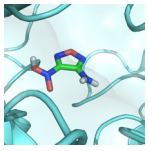
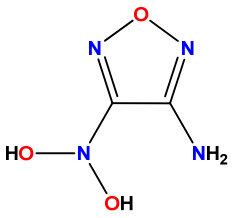
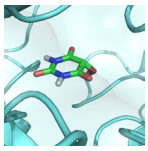
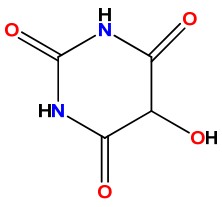
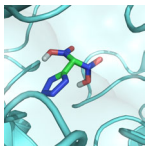
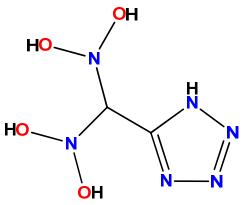
Compound dockposes	2D structure	MLogP	H-bond donor	H-bond acceptor	AutoDock score (kcal/mol)	Binding affinity
 MW 206.05		-2.832	7	7	-8.43	660.42 nM
 MW 206.03		-2.832	5	7	-8.34	773.19 nM
 MW 178.1		0.636	0	8	-8.24	908.45 nM
 MW 196.20		1.253	2	4	-8.23	921.54 nM
 MW 150.22		-2.055	4	3	-8.09	1.17 µM

Table 7.1 contd.

b

Compound dockposes	2D structure	MLogP	H-bond donor	H-bond acceptor	Vina score (kcal/mol)
 MW 160.08		-2.348	4	7	-8
 MW 146.08		-1.177	4	7	-7.9
 MW 132.08		0.173	4	7	-7.9
 MW 144.09		-2.401	3	6	-7.9
 MW 178.11		0.596	5	10	-7.7

ii) COmbining Docking And Similarity Search (CODASS)

The in-house high-throughput virtual screening program, LIDAEUS is a screening approach based on the generation of site points (categorised into hydrogen bond donor, hydrogen bond acceptor or hydrophobic) around the pocket of a target protein (Taylor et al. 2008). Molecules selected from a database are docked individually and matched to the generated site points in the binding pocket based on the preferred interactions. The predicted interactions are therefore scored by two methods: i) force field-based energy function combining both van der Waals and hydrogen bond energies and ii) pose interaction profile which is specific to a particular type of interactions (Taylor et al. 2008), and expressed by the energy scores (kcal/mol). While a number of publicly-accessed small molecule databases are available for data-mining, EDULISS, the in-house relational small molecule database provides a collection of over four million commercially available compounds from 28 different suppliers (Hsin et al. 2011). This allows high-throughput searching capabilities, and thus provides many possibilities for specific protein-ligand interactions to be explored.

In order to obtain comparable and accurate results from virtual screens, a group of screening programs were combined to provide a consensus prediction of possible molecular interactions. This involved a cycle of parallelised fast flexible docking and similarity search algorithms comprising AutoDock, AutoDock Vina and LIDAEUS, which were employed against the EDULISS database. The analysis was performed by Dr. Douglas Houston, using *Lm*iPGAM as the target molecule and 3PGA as the original ligand, in the presence of Zn^{2+} atoms. The scores from all three programs were compared and ranked based on "consensus score", which indicates the agreement of each algorithm, where the lower the number, the better. The full results are shown in the Appendix II section, while the top five hits, ranked based on the "consensus scores" are illustrated in Table 7.2. Selected candidates from the screen, together with the previous AutoDock and AutoDock Vina results, were pooled and compared before compounds were purchased for inhibition analysis. The basis of the selection of compounds is discussed in section 7.5.

Table 7.2 The top five binding candidates from the combination of docking and similarity search experiments.

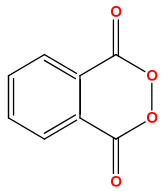
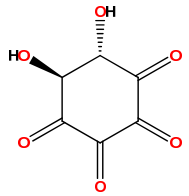
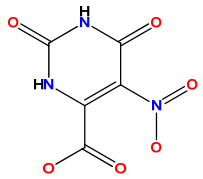
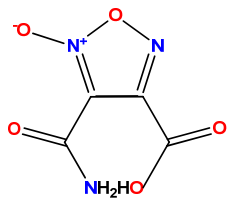
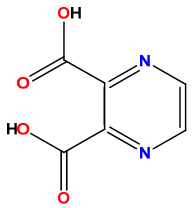
Compound	Molecular weight	MLog P	H-bond donor	H-bond acceptor	AutoDockScore (kcal/mol)	Vina score (kcal/mol)	Xscore	Drug score	Consensus score
	164.12	1.804	0	4	-9.14	-6.2	5.74	-314389	164.875
	172.09	-3.328	2	6	-8.2	-6.4	5.76	-302688	190.875
	200.09	-1.825	2	9	-9.48	-5.3	5.96	-288453	207.75

Table 7.2 contd.

	173.1	-1.477	3	8	-9.9	-7.1	5.72	-282963	208.125
	168.1	-1.254	2	6	-9.13	-5.6	5.73	-312739	210.875

7.4 Quantitative high-throughput screening strategy

As a complement to the *in silico* virtual screening approach, quantitative high-throughput screening (qHTS) for *Lm*iPGAM is currently being conducted by the NIH research group. The methodology relies on large-scale automated titration-based screening against a large number of compounds from a comprehensive chemical library (Inglese et al. 2006; Inglese et al. 2007a; Inglese et al 2007b). The major advantage of this approach is the ability to screen compounds at multiple concentrations, thus minimising false positive and false negative hits, as well as extensive follow-up testing (Inglese et al. 2006).

With a microtiter-plate based platform, a molecular target end-point assay for *Lm*iPGAM was performed in the presence of the biochemical components involved in the reaction. Figure 7.7 depicts both the forward and reverse reactions involved in assaying *Lm*iPGAM activity, where the enzyme luciferase is utilised for the generation of luminescence as a detection method. In this case, the Kinase-Glo® Luminescent Kinase assay (Promega) method was utilised to quantify the amount of ATP remaining in the sample, after a kinase reaction. As discussed in chapter 6, the activity of iPGAM is not readily measured, hence coupling enzymes are required to be included. For the forward reaction, both enolase (ENO) and PYK are involved, whereas in the reverse reaction, both GAPDH and PGK are essential for the biochemical reaction. It is noteworthy that the generated luminescent signal is inversely proportional to the amount of kinase activity, and correlated with the production of ATP in the reaction (http://www.promega.com/products/cell-signaling/protein--kinases-and-kinase--assays/kinase_glo-luminescent-kinase assays /).

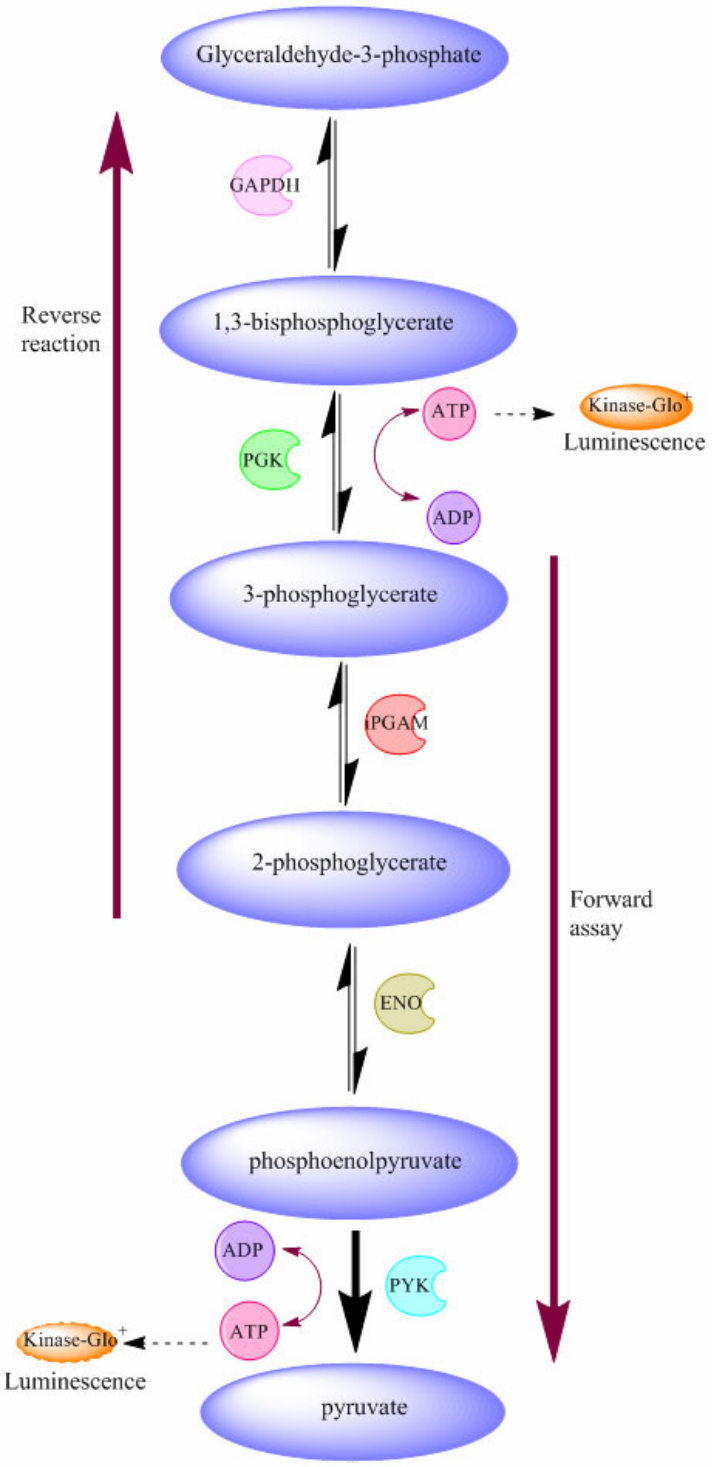


Figure 7.7 A schematic diagram showing the forward and reverse reactions for assaying *Lm*iPGAM activity by luminescence detection.

It is typical in high-throughput screening experiments for an analysis of the concentration response curve (CRC) to be conducted and classified. This is vital to identify any activators or inhibitors from the primary screen with desirable potencies for subsequent investigation of structure-activity relationships (Inglese et al. 2006). The CRCs can be divided into four main categories based on the quality of the curve fit (r^2), the efficacy (magnitude of response) and the number of asymptotes to the calculated curve: i) class 1 with an r^2 value > 0.9, with both lower and upper asymptotes shown in the complete CRCs, ii) class 2 with incomplete CRCs lacking lower asymptote and r^2 value > 0.9, iii) class 3 with compounds only exhibiting activity at the highest concentration tested, and iv) class 4 which contains compounds without CRC and thus classified as inactives (Inglese et al. 2006) (Figure 7.8 and Table 7.3).

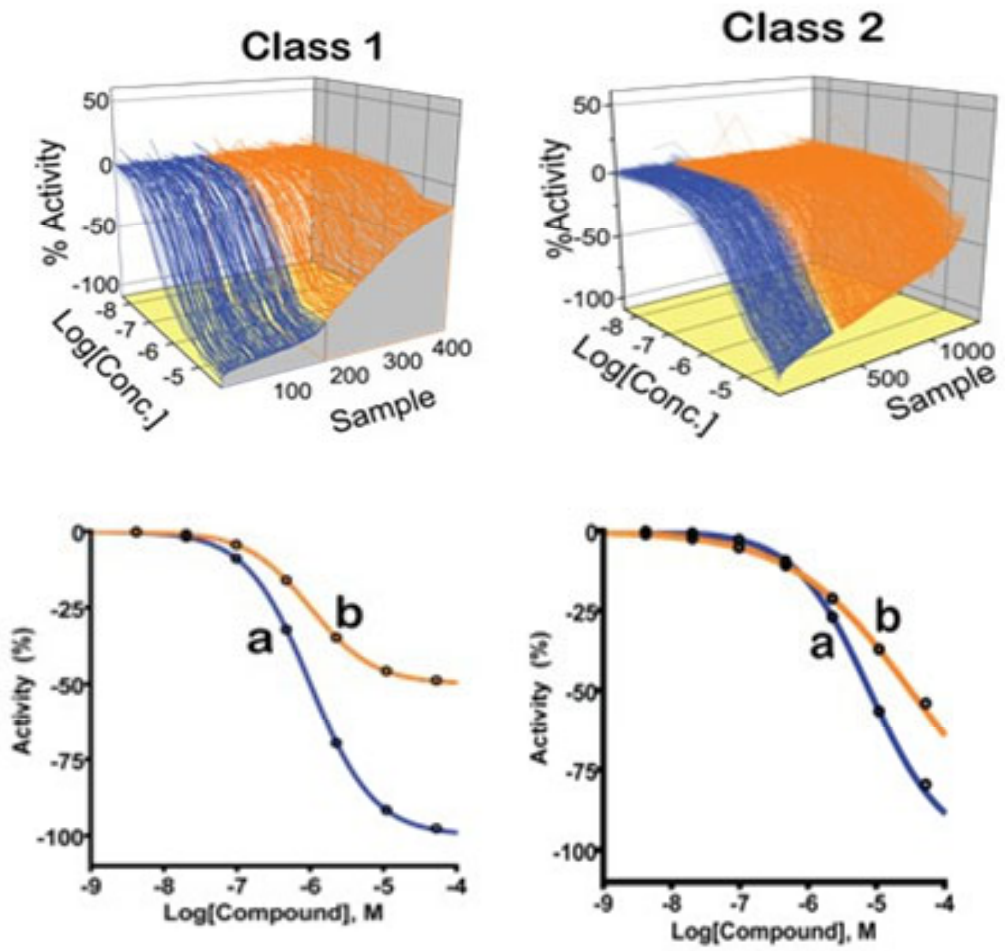


Figure 7.8 Concentration response curves (CRCs) are classified into four types. Examples of curves fitting the following classification criteria: Class 1 curves display two asymptotes, an inflection point, and $r^2 \geq 0.9$; subclasses 1a vs. 1b are differentiated by full (>80%) vs. partial (≤ 80%) response. Class 2 curves display a single left-hand asymptote and inflection point; subclasses 2a and 2b are differentiated by a max response and r^2, >80% and >0.9 or <80% and <0.9, respectively. Class 3 curves have a single left-hand asymptote, no inflection point, and a response >3SD the mean activity of the sample field. Class 4 defines those samples showing no activity across the concentration range. The figure and the table were obtained from a presentation by Dr. Linda Gilmore.

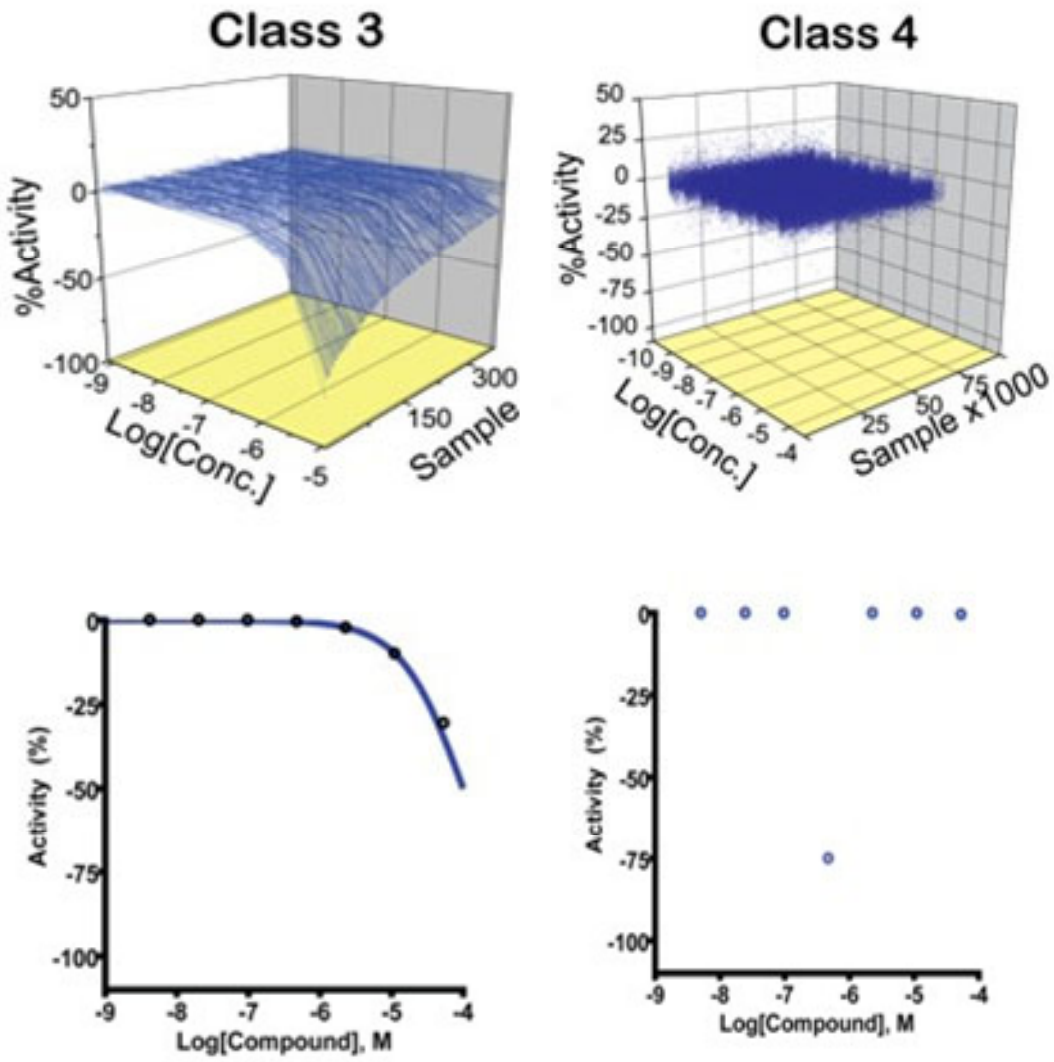


Figure 7.8 contd

Table 7.3 The main characteristics of the four different classes of the CRCs, as described above in the legend for Figure 7.8.

Description	**1 Complete response** **2 Partial response**	**Incomplete curve**	**'Single point' activity**	**Inactive**
Assymptotes	2	1	1	0
Inflection	Yes	Yes	No	No
Efficacy magnitude of response	1 >80% 2 ≤80%	1 >80% 2 <80%	>3 SD from mean activity at highest concentration	NA
r^2	≥0.9	1 >0.9 2 <0.9	NA	NA

The primary qHTS screen for *Lm*iPGAM was conducted initially with the Library Of Pharmacologically Active Compounds (LOPAC®), which contains pharmaceutically diverse compounds with known drug or drug-like activity (Sigma-Aldrich). The results are shown in Figure 7.9, where the hit rate of iPGAM was very low, probably as a consequence of its poor accessibility to the active site. Only 0.08% from the hits was typified as Class 1 and another 0.47% was classified as Class 2 inhibitors. Out of the hits, seven active compounds were obtained, as shown in Figure 7.10, where only three compounds were suitable for further testing in secondary screens. The rest of the hits were a known luciferase inhibitor, a common false positive, or exhibited some cross activity (Dr. Linda Gilmore, personal communication). The three hit compounds with IC_{50} values between 13-26 μM were evaluated by secondary screening in the Edinburgh Structural Biochemistry group laboratory.

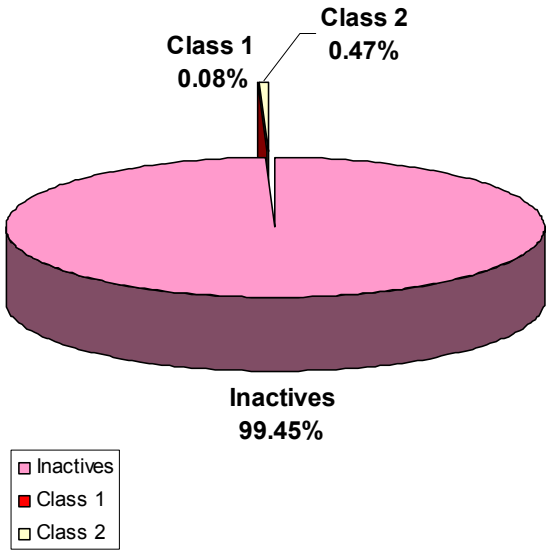


Figure 7.9 The percentage of *Lmi*PGAM LOPAC® hits classified as Class 1 and Class 2 inhibitors (0.08% and 0.47%), respectively. The remaining 99.45% compounds were classified as inactives. The figure was adapted from a presentation by Dr. Linda Gilmore.

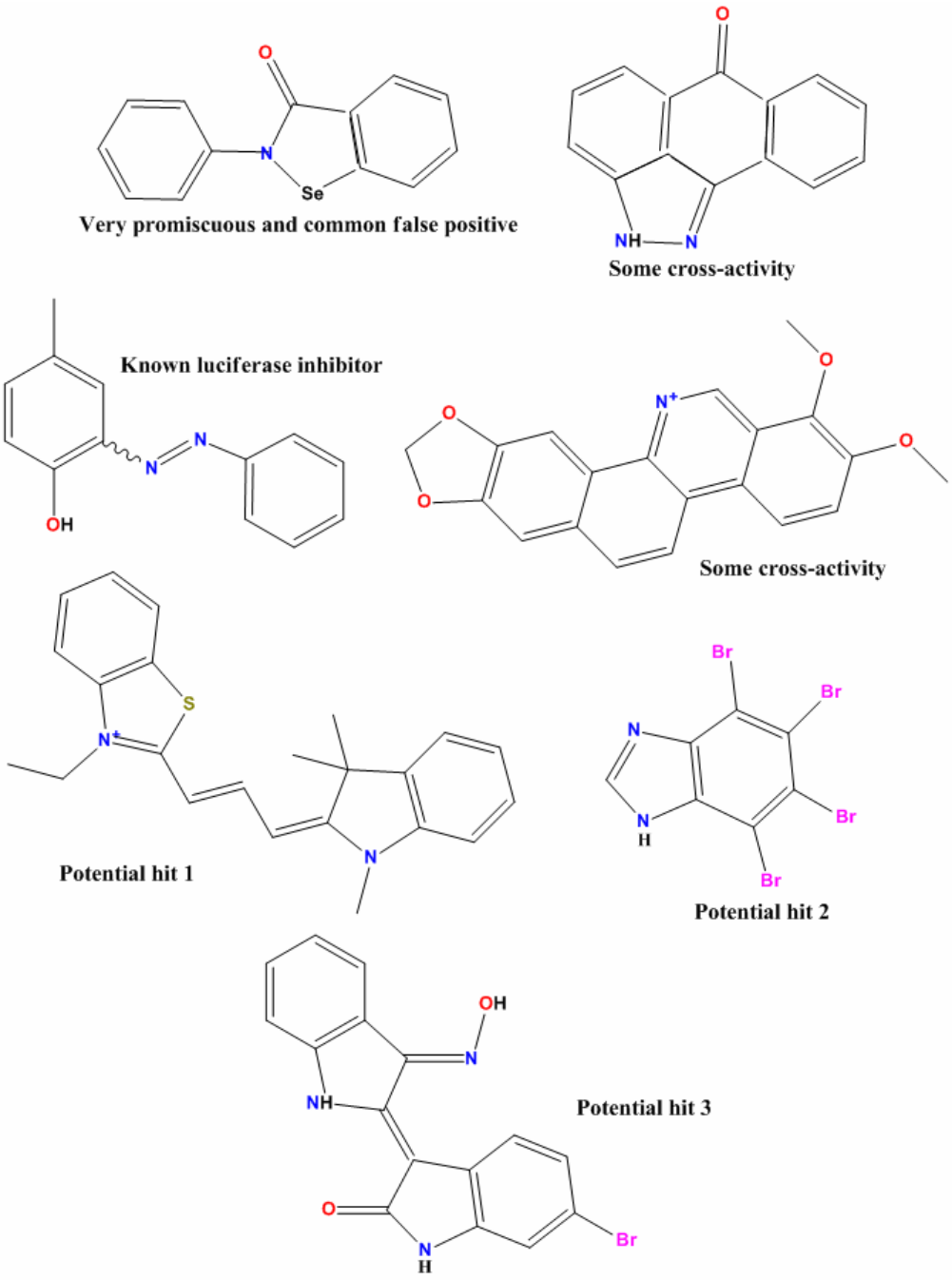


Figure 7.10 Seven active compounds from the initial LOPAC® screen. Only three compounds were considered as potential hits, whereas the rest were either known as a luciferase inhibitor, a common false positive or exhibited some cross-activity. The figure was adapted from a presentation by Dr. Linda Gilmore.

7.5 Preliminary inhibition analysis with novel compounds

It has always been a challenge to find appropriate drug candidates which exhibit not only possibilities to bind to the target protein, but also show significant effects on the protein's activity. While the *in silico* virtual screens provide successful prediction of the binding affinities of the compounds and the target protein, and the qHTS enables the assessment of binding affinities and initial inhibitory effects on the target molecule, secondary analyses are also necessary to confirm the specificity of the inhibitors. It is therefore important to perform an additional inhibitor screening with the standard enzyme assays that are routinely implemented in the laboratory. In this case, both the continuous and discontinuous coupling enzyme assay approaches (see chapters 2 and 6) were employed to screen the potential compounds.

Based on the initial LOPAC® screen, three compounds emerged as potential candidates to inhibit *Lm*iPGAM activity. Compound C1 and C3 (6-bromoindirubin-3′-oxime (BIO) (Table 7.4) resembled 'potential hit 3' in Figure 7.10, and was obtained from Enamine and Sigma-Aldrich, respectively. Compound C2 rabromobenzimidazole (TBBz) corresponds to 'potential hit 2' in Figure 7.10, and was from Sigma-Aldrich. In addition to these two selected compounds, a number of virtual screen 'hits' were also included for the activity assays. Four compounds (C4-C7) were selected from Enamine, while four others (C8-C11) were obtained from Interbioscreen. The compounds were chosen by Dr. Douglas Houston and were considered for the subsequent inhibitory analysis since they covered a range of chemical classes, were inexpensive and readily available. Some of these compounds contained specific chemical groups which may resemble the properties of the substrate and product, phosphate mimics or carboxylic acid groups. They were also chosen based on their solubility and binding energy. The structures and molecular properties of the selected compounds are shown in Table 7.4.

A number of amino acids with different functional groups were also included in the inhibition analysis, and were divided into two main classes: i) polar and hydrophilics (L-aspartate, L-glutamate, L-serine), and ii) non-polar or hydrophobics (L-valine, L-cysteine, L-leucine and L-isoleucine). L-isoleucine, which has

previously been shown to be an inhibitor of *Lm*iPGAM (Poonperm 2005) was included as a positive control. For the purpose of observing whether molecules in other metabolic pathways can also affect iPGAM activity, α-ketoglutarate, a key intermediate in the Krebs cycle, and malonate, a well-known competitive inhibitor of succinate were also included in the inhibition experiments (Table 7.5).

Table 7.4 Selected compounds from LOPAC® and virtual screens for the inhibition analysis. In C4, only the right hand molecule was involved in the docking experiment, while the left hand molecule was present in the compound, and most probably included as a counter ion. The presence of the left hand molecule may affect the molecular weight of the compound, thus also affecting the concentration calculations.

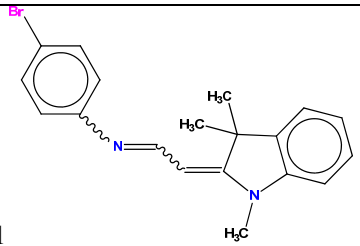
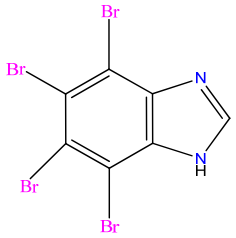
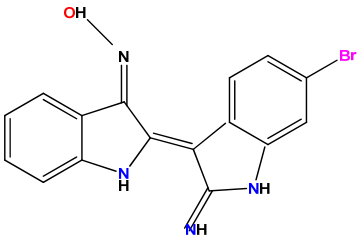
Molecule	ID number	Molecular weight	Supplier
 C1 Similar to NIH	T0501-7852	355.3	Enamine
 C2 Similar to NIH	TBBz T6951	433.7	Sigma-Aldrich
 C3 Similar to NIH	Sigma-Bio B1686	356.2	Sigma-Aldrich

Table 7.4 contd

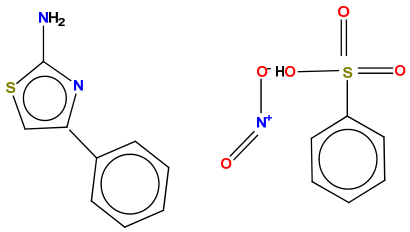
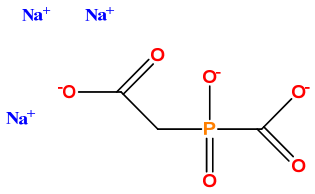
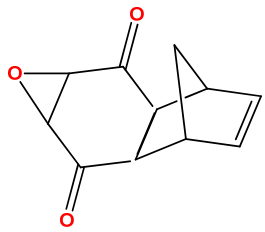
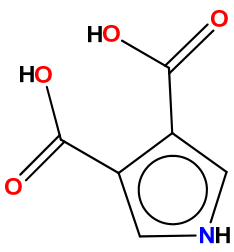
 C4 31DRH1-866-611	T5678776	379.4 (both molecules)	Enamine
 C5 31DRH1-475-257	T0500-7018	234.0	Enamine
 C6 31DRH1-475-336	T0500-7535	190.2	Enamine
 C7 31DRH2-114-014	T6570991	155.1	Enamine

Table 7.4 contd

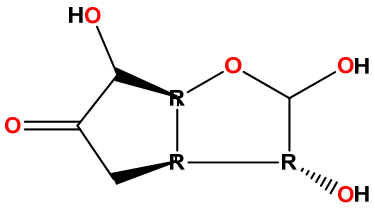
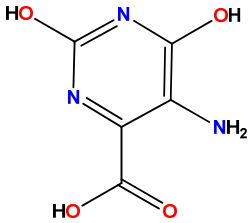
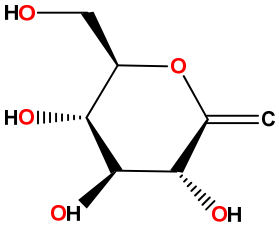
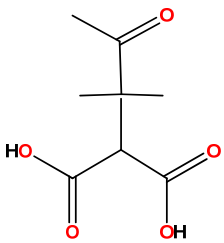
 C8 8DRH1-001-475	BB_NC-0551	176.1	Interbioscreen
 C9 8DRH1-001-018	BB_NC-0020	171.1	Interbioscreen
 C10 8DRH1-001-835	BB_NC-1065	178.1	Interbioscreen
 C11 8DRH1-001-266	BB_NC-0311	188.2	Interbioscreen

Table 7.5 Selected amino acids and metabolites for the inhibition analysis.All compounds were purchased from Sigma-Aldrich.

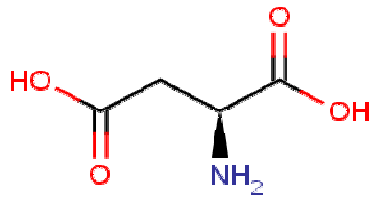
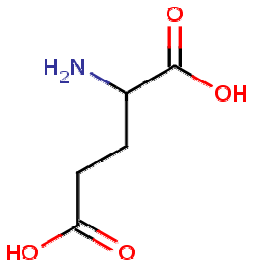
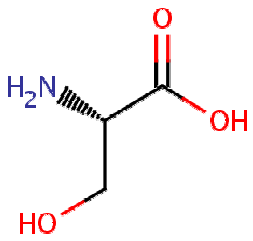
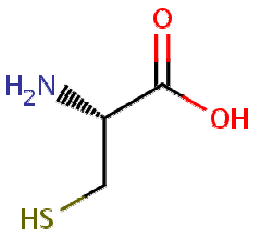
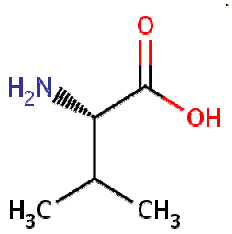
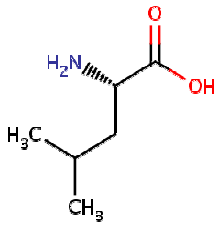
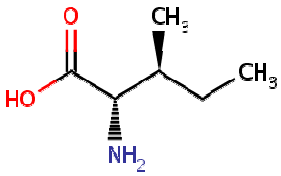
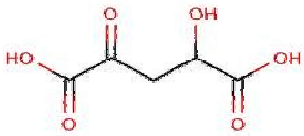
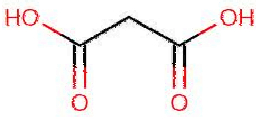
Molecule	Molecule name	Molecular weight
	L-aspartic acid	133.1
	L-glutamic acid	147.1
	L-serine	105.1
	L-cysteine	121.2

Table 7.5 contd.

	L-valine	117.1
	L-leucine	131.1
	L-isoleucine	131.1
	α- ketoglutarate	146.1
	malonate	104.1

Since dimethyl sulfoxide (DMSO) (Figure 7.11a) was utilised for the purpose of dissolving the compounds, it is therefore important to determine the tolerances of the coupling enzymes and *Lm*iPGAM towards the organosulfur compound. Figure 7.11b depicts the effect of DMSO on the coupling enzymes (ENO, PYK and LDH) and on *Lm*iPGAM, where the range of DMSO concentrations tested was from 1-10%. Enzyme samples in the absence of DMSO were included as negative controls. It is apparent from the results that *Lm*iPGAM is more susceptible to higher concentrations of DMSO, where substantial effects on the activity could be observed when the concentration exceeded 2%. The remaining activity decreased as the concentration of DMSO rose. Interestingly however, the coupling enzymes exhibited better tolerances to DMSO, which became more apparent when compared to iPGAM at the maximum concentrations tested (9% and 10% DMSO). This intriguingly suggested that *Lm*iPGAM is specifically inhibited by DMSO, and therefore the maximum final concentration of DMSO that was present in the sample during the inhibition analysis has been standardised to 1-2%.

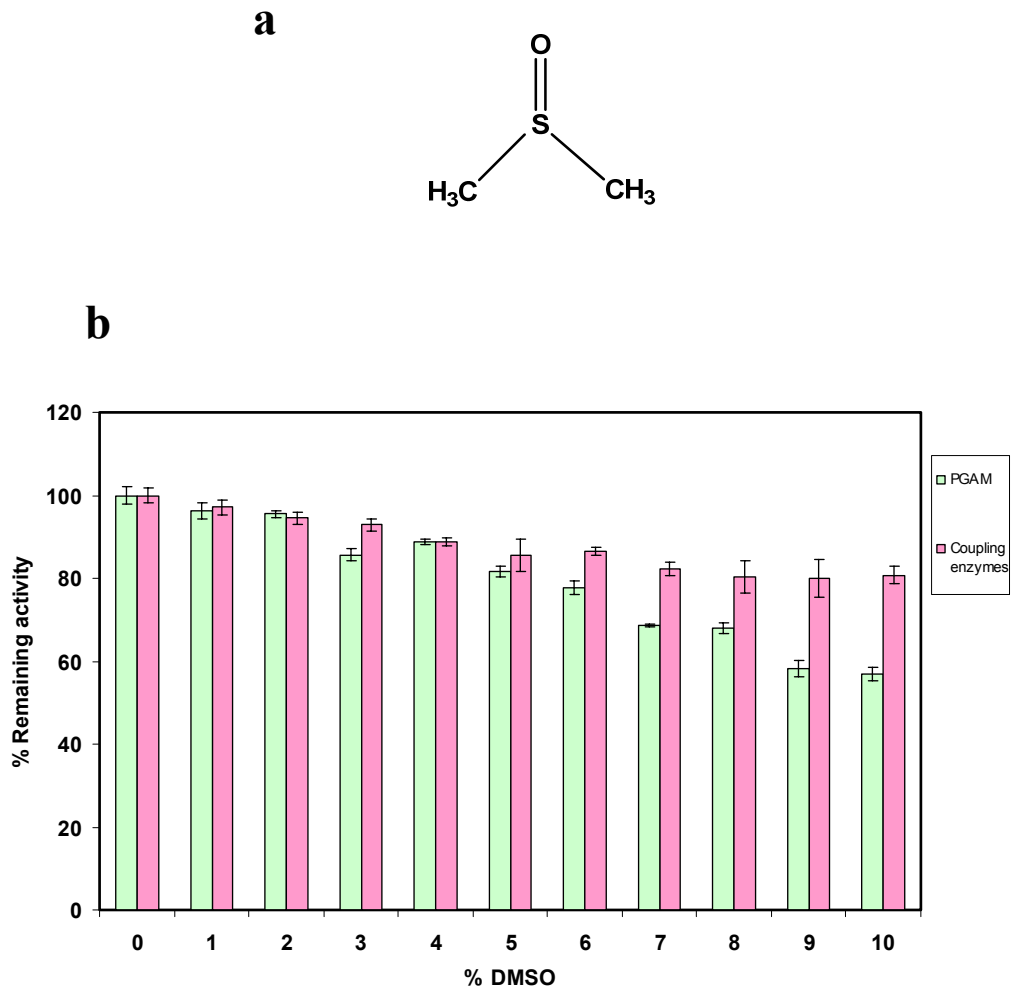


Figure 7.11 a) The chemical structure of DMSO and b) DMSO tolerance test for the coupling enzymes (ENO, PYK and LDH) and for *Lm*iPGAM.

The inhibition analysis was performed by incubating 0.162 µM P2 *Lm*iPGAM from the ion-exchange column with a final concentration of 1 mM Co^{2+} at 25°C for 1 h, before adding a final concentration of 1 mM compounds individually with further incubation for another hour. P2 was chosen for the inhibition analysis because it is in the open conformation, and this may provide better access to the active site. It is noteworthy that a relatively high concentration of buffer was chosen in the experiment, in order to minimise the effects of compounds to pH. The cobalt incubation was employed to activate the enzyme to its maximum activity (see chapter 5 and 6) before observing the effects of the compounds. The results are shown in Figure 7.12, where the inhibitor effects were tested on both *Lm*iPGAM and

the coupling enzymes in the continuous assay. The discontinuous assay (Figure 7.13) however, carried an advantage since the inhibitor concentrations would be 20-fold lower in the well containing the coupling enzymes. The methods for both analyses can be found in chapter 2 and specifically in chapter 6 for the discontinuous assay. The basis of the reaction was similar to the metal studies (see chapter 6), hence the effects on the coupling enzymes were minimised. In both cases, enzymes in the absence of the compounds were included as negative controls, and the analyses were conducted in duplicate. The samples were tested in the presence of 1% DMSO, except for C5 and C10, which were done in the presence of 2% DMSO. This is due to the insolubility of the compounds that required to be dissolved at a higher concentration of DMSO. Nevertheless, the effect of DMSO at 1% and 2% were insignificant, thus the results could be equally compared.

The efficacy of some inhibitors could be observed in the continuous assay (Figure 7.12), where the activity of *Lm*iPGAM was specifically inhibited by α-ketoglutarate, malonate, C1 and C2 from the LOPAC® screen, and some of the virtual screen hits: C4, C5, C7 and C11. The efficacy of some compounds however, was observed to be better than others by showing more than 50% inhibition, for example C7 (49% remaining activity) and C2 with 14% remaining activity, when compared to the control sample. C1, C4, C5 and C11 exerted an effect by causing remaining activity to be in the range of 87% to 96% respectively. C3 on the other hand, showed a tremendous effect not only on *Lm*iPGAM, but on the coupling enzymes as well, thus suggesting a non-specific inhibition. By contrast, C6, C8 and C10 did not show any specific inhibition of *Lm*iPGAM, and C9 peculiarly, did not show any signal from the absorbance measurement, and thus suggested to be a strong activator for the coupling enzymes. The Krebs cycle intermediate α-ketoglutarate exhibits partial inhibition with a remaining activity of 89%, while malonate did not cause a significant effect on *Lm*iPGAM. The rest of the compounds exhibit an activating effect and this was apparent in the samples containing amino acids, particularly L-aspartate and L-cysteine. L-isoleucine, which was previously reported to be an inhibitor however, did not exhibit any significant effect on *Lm*iPGAM activity.

By contrast, the outcome from the discontinuous assay (Figure 7.13) exhibits relatively small effects of the compounds on *Lm*iPGAM activity. With minimal interference on the coupling enzymes, the results (expressed as the concentrations of 2PGA after 15 min incubation) indicate that the activity of *Lm*iPGAM remained essentially the same, with the formation of 0.111 to 0.157 mM 2PGA in the tested samples, compared to the control in the absence of inhibitor (0.154 mM). By contrast to the continuous assay however, there seems to be no obvious activating effects of these compounds on *Lm*iPGAM, and malonate, C4 and C10 in particular, exhibited ~20% inhibition of *Lm*iPGAM. This thus suggested that the compounds may prevent iPGAM from reaching its equilibrium when compared to the control sample, which reached equilibrium within the incubation period. The trend of inhibition may well be due to the specific chemical properties of the compounds, such as the presence of carboxylic acid group in most of the compounds from virtual screens and in the amino acids, or specific functional groups that may interact with the residues in the active site. The different results obtained from both analyses may indicate the involvement of the coupling enzymes in affecting iPGAM activities in the continuous assay. Further trials are required to be conducted to provide significant understanding of the effects of the inhibitors on *Lm*iPGAM activity. Apart from that, it is noteworthy that a piperazine analogue: 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid has been crystallised with *Ba*iPGAM strain Sterne (Chang et al. 2011, PDB ID 3R7A) (Figure 7.14a and b). Piperazine, which is a six-membered ring containing two opposing nitrogen atoms has been widely used as anthelmintic (Holden-Dye and Walker 2007) and antileishmanial agents (Mayence et al. 2004). The effect of piperazine derivative on *Bs*iPGAM is yet under development, but may well be a potential candidate that explicitly affects the transferase domain for future inhibition studies.

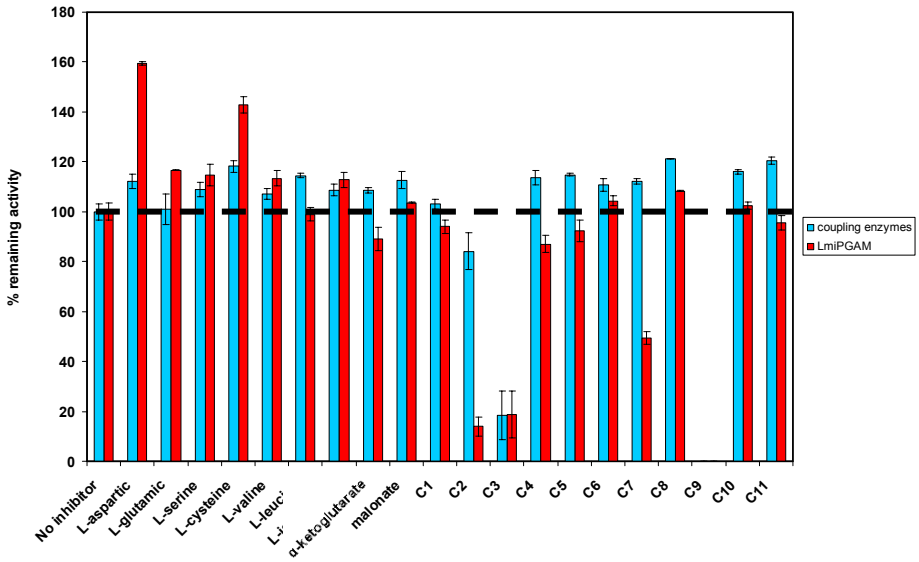


Figure 7.12 Inhibition analysis (see Tables 7.4 and 7.5) for the coupling enzymes and *Lm*iPGAM. 2PGA or 3PGA were used as the starting reagents for the coupling enzymes (ENO, PYK and LDH) or *Lm*iPGAM, respectively. The assay was carried out in a continuous fashion with 1mM of each compound.

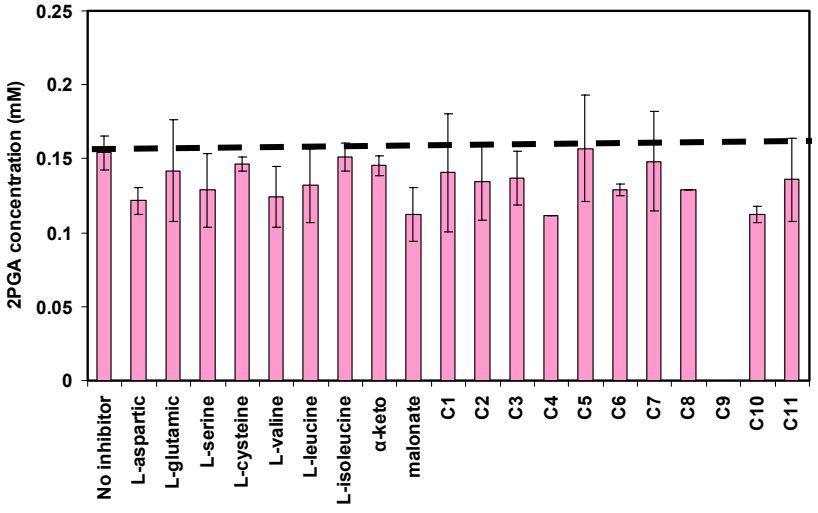


Figure 7.13 Inhibition analysis (see Tables 7.4 and 7.5) using the discontinuous assay. The results are expressed as the concentration of 2PGA after 15 min incubation of *Lm*iPGAM with 1mM of each compound.

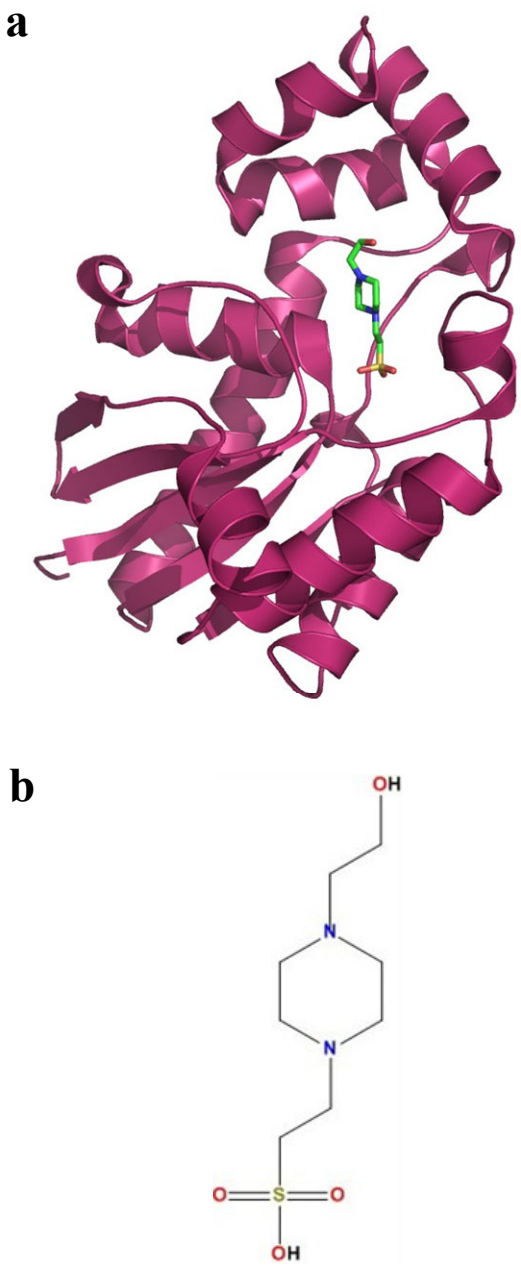


Figure 7.14 a) The crystal structure of the transferase domain of *Ba*iPGAM strain Sterne (Chang et al. 2011, PDB ID 3R7A) complexed with 4-(2-hydroxyethyl)-1-piperazine ethanesulfonic acid, and b) Its chemical structure.

7.6 Conclusion

Two different approaches for *Lm*iPGAM drug development have been employed. The qHTS which is currently being conducted in the NIH, Bethesda, US primarily yielded a low number of hit candidates from the small-scale (LOPAC®) screen, and this is potentially due to the small buried active site between the two domains. Nevertheless, there were three promising inhibitor candidates with relatively low IC_{50} values, and two of them were selected for secondary screening. The second approach was conducted via the *in silico* virtual screening method, and was primarily employed by the ligand-based strategy. This is a direct consequence of the inaccessibility of the active site, which required the analysis to be conducted based on the structural and molecular properties of the substrate/product. A series of structure-based screens through a similarity search, as well as multiple docking and scoring algorithms were later employed. While virtual screens provide a higher number of potential binding candidates for the protein, high-throughput screens have the advantage of primary inhibition analysis prior to a secondary screen. Nevertheless, the outcome from both approaches was compared and nine compounds with distinct molecular properties from different chemical classes were selected for secondary screens.

Subsequent secondary analysis using the continuous and discontinuous coupling enzyme assays were performed for 11 selected compounds. Seven amino acids with different molecular properties, together with α-ketoglutarate and malonate were also included in the inhibition analysis, giving a total of 20 compounds tested. Through the continuous assay, two potential inhibitors were obtained, with less than 50% remaining activities (C2 NIH with 14% remaining activity) and C7 (49% remaining activity). The discontinuous assay however, showed only modest effects of the inhibitors on the activity of *Lm*iPGAM. The 2PGA concentrations were in the range of 0.111 to 0.157 mM, as compared to the 0.154 mM 2PGA in the control sample at equilibrium. The differences between the two experimental approaches require further analysis in future trials.

Chapter 8:

Concluding Remarks and Future Directions

8.1 General conclusions

An understanding of protein biochemistry is fundamental in drug design and discovery. Without comprehending the nature of the target molecule, the efficacy of the chosen therapeutic approach may well be illusory. While current molecular approaches tend to use *in vitro* technology and have been successful in mimicking *in vivo* cellular conditions, there are also experimental flaws that contribute to data inaccuracy. In addressing this issue, experiments have to be carefully designed with a preference for biologically relevant conditions which have to be taken into account. The aims of this study, as discussed earlier in chapter 1, are focussed on two main biochemical and structural aspects of the enzyme phosphoglycerate mutase from trypanosomatids, and concern: i) The biologically relevant metal ions that support the enzyme activity and ii) Characterisation of the enzymes to obtain druggable states, which are important for designing and discovering potential inhibitors. This study is therefore performed on the basis of understanding the behaviour of the enzyme under conditions that closely resemble the *in vivo* situation, when the parasite is infecting a human.

Trypanosomatid iPGAMs are drug targets. Trypanosomatid iPGAMs, particularly from *T. brucei* have been genetically validated as potential drug targets. This has been shown through RNAi experiments (Albert et al. 2005), and the observation that the pathogenic bloodstream-form particularly depends on glycolysis for a source of ATP. *Lmi*PGAM, which is the corresponding enzyme in *L. mexicana* is also a subject of interest in kinetoplastid chemotherapy. The latter parasite however, possesses a very different biology compared to the former one, especially once it is inside the mammalian hosts. While the *T. brucei* bloodstream form lives freely in the human blood, the *L. mexicana* amastigote form lives within phagolysosomes of macrophages (Barrett et al. 1999), and thus possesses a more complicated life mechanism. Nevertheless, the high sequence similarity observed between the two iPGAMs and the fact that both enzymes possess unrelated structural properties with

the human dPGAM further augmented the values of these enzymes as potential drug targets. iPGAM has also been verified to be a cytosolic enzyme (Verlinde et al. 2001; Guerra et al. 2004; Gualdrón-López et al. 2011) and this also provides an advantage in better access for the drugs, rather than having to go through the glycosomal membrane to reach the target proteins.

The iPGAM metal requirements. iPGAMs are classified as members of the alkaline phosphatase superfamily, which is characterised by enzyme metal dependency. The explicit requirement of Co^{2+} for structural and functional integrity of trypanosomatid iPGAMs has been widely elucidated, but remains controversial since the known cellular concentrations of Co^{2+} in trypanosomatids and other eukaryotes are insufficient *in vivo*. To address this problem, a new purification approach for the bacterially expressed enzyme has been designed in this study, by substituting Co^{2+} with Ni^{2+} for the initial IMAC purification step. Two additional purification stages were subsequently involved, which are the desalting and ion-exchange chromatography. The purified samples which were used for metal studies were therefore estimated to be free from Co^{2+} ions, and were also subjected to various structural and biochemical analyses.

The different conformations and oligomeric states of LmiPGAM. Ion-exchange chromatography has successfully separated the bacterially expressed *Lm*iPGAM into different conformational and oligomeric states, which have been characterised by a series of biophysical analyses, notably SEC-MALS. The enzyme was mainly observed in four states: the closed-form monomeric (P1), open-form monomeric (P2), and closed and open dimeric forms. The presence of higher order oligomers was also observed in the most unstable form of the enzyme (P3). Moreover, it was observed that in the presence of Co^{2+}, the enzyme was pushed towards the closed-form conformation, which is the active state of the enzyme. While this observation was noted *in vitro*, this may not be the case *in vivo*, where Co^{2+} was not present in access. Hence, it is postulated that the majority of the native enzyme may exist in the inactive open conformation. This may well be related to the low specific activity observed in the native enzyme (Guerra et al. 2004; Albert et al. 2005; Paul Michels, unpublished data), compared to the corresponding bacterially expressed one

(Poonperm 2005). The different types of analyses which were elucidated in this study also provide an understanding of the individual characteristics of each peak from the ion-exchange chromatography, which strikingly differ in terms of the shape and size of the molecules, the pace of catalytic function and metal content. Moreover, by comparing the structural aspects of the bacterial and trypanosomatid iPGAMs, it is very likely that these enzymes can be grouped into two different iPGAM families. This is observed by the difference in metal-mechanisms, which can be related to a larger open conformation of bacterial iPGAM compared to the trypanosomatid iPGAM. The former utilises both metal sites during catalysis, while the latter performs the catalytic function with a one-metal mechanism.

The hyperactivation of iPGAM with Co^{2+}. The activity of trypanosomatid iPGAM *in vitro* was shown to be maximally active in the presence of Co^{2+}. The *in vivo* activity however, is one of the lowest among the glycolytic enzymes. ICP-OES and ICP-MS analysis demonstrated that Co^{2+} is essentially undetected in both the cytosol of *T. brucei*, and the bacterially expressed *Lm*iPGAM. It is therefore concluded that the enzyme experiences hyperactivation in the presence of excess Co^{2+} *in vitro*.

Mg^{2+} and Zn^{2+} may act as iPGAM regulators in vivo. By contrast, Mg^{2+} and Zn^{2+} were found to be in excess *in vivo* with regard to iPGAM metal sites. To further investigate the biologically relevant metals for iPGAM, the activity needed to be assayed in the presence of Mg^{2+} and Zn^{2+}, together with a range of other divalent metal ions. Unfortunately, iPGAM activity cannot be readily measured in the absence of coupling enzymes in the continuous assay, and some of these enzymes, particularly ENO, are inhibited by higher concentrations of metals. To overcome this problem, a multimode-plate reader discontinuous assay has successfully been developed, with minimal metal interference on the coupling enzymes. The results revealed that of the biologically relevant metals, only Mg^{2+} can support the activity of iPGAM, although less than 50% in the presence of Co^{2+}. Zn^{2+}, on the other hand, exhibits an inhibitory effect on iPGAM. It is therefore suggested that Zn^{2+} may contribute to the lower activity measured *in vivo* in the absence of excess Co^{2+}, and that iPGAM activity in the cell is balanced by the relative concentrations of Mg^{2+} and Zn^{2+}. The discontinuous assay also revealed that the substrate:product ratio at

equilibrium in solution is 12:1, which is distinguishable from the 1:1 ratio observed from the crystal structure. The tighter binding of 2PGA may well be the reason behind this observation.

Preliminary screening approaches for iPGAM. For drug design, two screening approaches were developed to find potential inhibitors for *Lm*iPGAM. The first approach is performed by a collection of *in silico* virtual screening programs, comprising a similarity search, and subsequent docking and scoring experiments. The second approach is the high-throughput screening which is currently being conducted in the NIH, Bethesda. By compiling the outcomes from both screens, a secondary screen was conducted using both the continuous and discontinuous assays. Out of a collection of 11 compounds, two showed explicit inhibition of iPGAM at 1mM, with 14% and 49% remaining activity by using the continuous assay. The results from the discontinuous assay however, exhibited only slight inhibition of these compounds on iPGAM, although tested at an identical concentration. Further inhibition analysis, such as medicinal chemistry and/or structure-activity relationship (SAR) thus may well be required to validate these observations.

8.2 Significant contributions

This study has contributed new ideas and has substantially improved the existing knowledge concerning the biochemical and structural aspects of trypanosomatid iPGAMs. These aspects are discussed in more detail below:

i) Two different families of iPGAM are distinguishable by distinct metal-mechanisms and structural properties

The sequence alignment of iPGAM across a fairly wide range of organisms strikingly showed that the enzyme could be divided into two different families. The trypanosomatid and plant iPGAMs are grouped together, while the bacterial, nematode, algal and fungal iPGAMs may well be unified under a different family. The distinction is highlighted by the metal-mechanisms, where iPGAMs from the first family are likely to utilise one metal site during catalysis, whereas members from the second family utilise both sites for catalytic function. In addition,

trypanosomatid iPGAMs possess a smaller open conformation, while the bacterial iPGAMs exhibit substantial movement of the two domains, resulting in a larger apoenzyme conformation.

ii) The separation of different forms and oligomeric states of iPGAM by ion-exchange chromatography

*Lm*iPGAM has successfully been separated into different forms and oligomeric states by ion-exchange chromatography. This observation can be the basis of obtaining the open-form structure, which is important for the purpose of drug design. The current closed-form enzyme provides tremendous structural information, but the poorly-accessible catalytic site caused difficulties in drug design implementation. It is also a striking observation that iPGAM, although it has been reported as a monomer previously, may also form a dimer and higher-ordered oligomers in solution. The interesting behaviour of the enzyme in solution also provides indications of the preferable conditions for crystallisation.

iii) The hyperactivation of iPGAM with Co^{2+} *in vitro* and low specific activity as observed *in vivo*

One of the main aims of this study was to elucidate the biologically relevant metal requirements of iPGAM. Metal analyses therefore successfully revealed that Co^{2+}, which was reported to support the trypanosomatid iPGAM activity, is essentially undetectable, both in the bacterially expressed enzyme, and also in the cytosol where the enzyme is located. This generates a new paradigm in understanding the nature of the enzyme in its native environment, where it is therefore suggested that iPGAM is hyperactivated by Co^{2+} *in vitro*. The insufficiency of Co^{2+} *in vivo* may result in low specific activity, which indicates that there are other biologically relevant metals, possibly Mg^{2+}, that assumed the role of Co^{2+} in the native environment. iPGAM is thus suggested to be one of the key players in controlling the parasite glycolytic flux by a balance of Mg^{2+} and Zn^{2+} as its regulators. In the study by Achcar et al. 2012, iPGAM was shown to maximally control the glycolytic flux, based on the control coefficients that have been analysed in the dynamic model. This may affect the

amount of ATP produced, and if this enzyme is specifically inhibited, the parasite, particularly the *T. brucei* bloodstream form, may experience difficulties in surviving in the mammalian hosts.

iv) The development of multimode-plate reader discontinuous assay

The continuous coupling enzyme assay is a means of assaying iPGAM activity by coupling the reaction with other enzymes involved in the glycolytic pathway. Unfortunately, the poor tolerances of the coupling enzymes, notably ENO, to high concentrations of metal ions made it impossible to assay iPGAM activity in the presence of a range of metal ions. It was therefore crucial to develop an assay system which could minimise the metal interference on the coupling enzymes. This study has successfully developed a discontinuous assay which can be performed by using a more economical plate-reader format. The outcome revealed that Mg^{2+}, which is the most abundant metal found *in vivo*, can support the activity of iPGAM, albeit at a lower level than Co^{2+}. Zn^{2+} on the other hand, inhibits the enzyme. It is also a significant observation that in the absence of the other coupling enzymes in the initial reaction, the substrate:product ratio is 12:1 at equilibrium in solution, and not 1:1 as observed in crystal structure.

v) The development of *in silico* virtual screens and high-throughput screens for *Lm*iPGAM

Both *in silico* virtual and high-throughput screens have been developed for *Lm*iPGAM. These analyses thus provide potential inhibitor candidates for the enzyme, which are significantly important for drug development. Two of the compounds showed substantial inhibition at 1 mM concentration, and are worth further characterisation in the future.

8.3 Future work

The findings from this study have paved the path to a better understanding of the enzyme structural and functional properties, which are important in relation to drug design and discovery. A number of potential research studies covering various aspects can be conducted, as discussed below:

i) Structural aspects

Biophysical characterisation provided substantial knowledge of the behaviour of iPGAM in solution. Based on these observations, it is indeed interesting to obtain the open-conformation crystal structure of *Lm*iPGAM. This can be done either by: i) identifying the preferable condition to obtain only the open conformation that is stable in solution and ii) performing crystallisation trials based on the conditions that have been standardised for *Tb*iPGAM. Thorough structural analysis could also be conducted for both the available closed-form *Lm*iPGAM and open-form *Tb*iPGAM, and residues involved in the hinge regions may be identified for explicit site-directed mutagenesis studies, as these residues may well be important in controlling the movement of the two domains. It is also interesting to observe the effect of divalent metals on the overall structure of iPGAM, which could be employed by fluorescence measurements/circular dichroisms analysis.

ii) Biochemical aspects

The metal analyses in this study have efficiently measured the concentration of the metals in both the purified protein and the cytosolic fractions. However, this type of analysis does not possess the ability to distinguish between the free and bound-metals in the samples. To further investigate this, it can be suggested that the macromolecules in the samples would have to be removed prior to the analysis, either by centrifugation with molecular weight cutoffs, or precipitation with trichloroacetic acid (TCA), where only the free metals can be detected afterwards. This may provide information on the availability of free metals that may bind to iPGAM in the native environment. More interestingly, better understanding of iPGAM metal requirement can be obtained by isolation of the native enzyme from the parasite, where subsequent structural and biochemical analyses can be performed to analyse the metal requirements. The low specific activity of iPGAM, especially in the bloodstream-form *T. brucei* has previously been observed, but a more thorough observation may be obtained by western blot analysis, where the concentration of iPGAM in both forms of parasites at different life stages can be compared. In addition, the specific activities of iPGAM in the different life stages of *T. brucei* and

L. mexicana would also be important to be measured, in conjunction with the western blot analysis.

iii) Drug design and discovery

Currently, *Lm*iPGAM is one of the target molecules for high-throughput screens in the NIH, Bethesda, USA. The outcome from this analysis can be tested by secondary screens, where various biophysical and structural analyses can be conducted, for example kinetic studies using the enzymatic assays and thermal stability test. Compounds which exhibit substantial potency may well be analysed through specific protein-ligand interaction studies, for example surface plasma resonance (SPR) and isothermal titration calorimetry (ITC). The results from the inhibition analyses performed in this study may also be characterised by subsequent medicinal chemistry analysis such as SAR, as well as analysing the effects of related compounds that have been modified by changing the chemical groups towards the activity of the enzyme. It will also be of particular interest to carry out crystallisation experiments to solve the crystal structure of iPGAM with the inhibitors bound to the active site or the hinge region. It is also interesting to consider inhibitors of other members of the alkaline phosphatase superfamily to design potential inhibitors for iPGAM. The compounds from virtual screens in this study which have not yet been investigated as potential inhibitors could also be analysed in similar further secondary analyses.

References

Abyzov, A., Errami, M., Leslin, C.M., and Ilyin, V.A. (2005). Friend, an integrated analytical front-end application for bioinformatics. *Bioinformatics* 21(18): 3677-3678.

Achcar, F., Kerkhoven, E.J., The SilicoTryp Consortium, Bakker, B.M., Barrett, M.P., and Breitling, R. (2012). Dynamic modelling under uncertainty: The case of Trypanosoma brucei energy metabolism. *PLoS Comp. Biol.* 8(1): e1002352.

Adie, J.E. (2010). Structure-based drug design of 11β-hydroxysteroid dehydrogenase type 1 inhibitors. PhD thesis. Institute of Structural and Molecular Biology, School of Biological Sciences. University of Edinburgh, Edinburgh.

Agresti, J. J., Antipov, E., Abate, A.R., Ahn, K., Rowat, A.C., Baret, J-C., Marquez, M., Klibanov, A.M., Griffiths, A.D., and Weitz, D.A. (2010). Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *PNAS* 107(9): 4004-4009.

Ainsworth, S., and Macfarlane, N. (1975). Activation and inhibition of rabbit muscle pyruvate kinase by transition metal ions. *Biochem. J.* 145: 63-71.

Albert, M.A., Haanstra, J.R., Hannaert, V., Roy, J.V., Opperdoes, F.R., Bakker, B.M., and Michels, P.A.M. (2005). Experimental and in silico analyses of glycolytic flux control in bloodstream form Trypanosoma brucei. *J. Biol.Chem.* 280 (31): 28306-28315.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2008). Molecular biology of the cell, 5th edition. Garland Science, New York and Abingdon.

Albery, W.J., and Knowles, J.R. (1976). Evolution of enzyme function and the development of catalytic efficiency. *Biochemistry* 15(25): 5631-5640.

Ali, M.H., and Imperiali, B. (2005). Protein oligomerization: How and why. *Bioorg. Med. Chem.* 13: 5013-5020.

Amato, V.S., Tuon, F.F., Bacha, H.A., Neto, V.A., and Nicodemo, A.C. (2008). Mucosal leishmaniasis. Current scenario and prospects for treatment. *Acta Trop.* 105: 1-9.

Anderson, R.A., Bosron, W.F., Kennedy, F.S., and Vallee, B.L. (1975). Role of magnesium in Escherichia coli alkaline phosphatase. *PNAS* 72(8): 2989-2993.

Andersson, M., Malmendal, A., Linse, S., Ivarsson, I., Forsen, S., and Svensson, L.A. (1997). Structural basis for the negative allostery between Ca^{2+}- and Mg^{2+}-binding in the intracellular Ca^{2+}-receptor calbindin D_{9k}. *Protein Sci.* 6: 1139-1147.

Andersson, C.E., and Mowbray, S.L (2002). Activation of ribokinase by monovalent cations. *J. Mol. Biol.* 315: 409-419.

Avilán, L., Gualdrón-López, M., Quiñones, W., González-González, L., Hannaert,V., Michels, P.A.M., and Concepción, J-L. (2011). Enolase: A key player in the metabolism and a probable virulence factor of trypanosomatid parasites-perspectives for its use as a therapeutic target. *Enzyme Res.* 2011: 932549.

Bakker, B.M., Michels, P.A.M., Opperdoes, F.R., and Westerhoff, H.V. (1997). Glycolysis in bloodstream form Trypanosoma brucei can be understood in terms of the kinetics of the glycolytic enzymes. *J. Biol. Chem.* 272 (6): 3207-3215.

Bakker, B.M., Westerhoff, H.V., Opperdoes, F.R., and Michels, P.A.M. (2000). Metabolic control analysis of glycolysis in trypanosomes as an approach to improve selectivity and effectiveness of drugs. *Mol. Biochem. Parasitol.* 106: 1-10.

Baneyx, F. (1999). Recombinant protein expression in Escherichia coli. *Curr. Opin. Biotechnol.* 10(5): 411-421.

Barrett, M. P., Mottram, J.C., and Coombs, G.H. (1999). Recent advances in identifying and validating drug targets in trypanosomes and leishmanias. *Trends Microbiol.* 7(2): 82-87.

Barrett, M.P., Burchmore, R.J.S., Stich, A., Lazzari, J.O., Frasch, A.C., Cazzulo, J.J., and Krishna, S. (2003). The trypanosomiases. *Lancet* 362: 1469-1480.

Behnke, W.D., and Vallee, B.L. (1976). The spectrum of cobalt bovine procarboxypeptidase A, an index of catalytic function. *PNAS* 69(9): 2442-2445.

Bertini, I., Gary, H.B., Lippard, S.J., and Valentine, J.S. (1994). Bioinorganic chemistry. University Science Books, Mill Valley, C.A.

Black, C.B., and Cowan, J.A. (1995). The biological chemistry of magnesium. VCH Publishers, New York.

Blaho, D.V., and Miranker., A.D. (2009). Delineating the conformational elements responsible for Cu(2+)-induced oligomerization of β-2 microglobulin. *Biochemistry* 48: 6610-6617.

Boehr, D.D., Nussinov, R., and Wright, P.E. (2009). The role of dynamic conformational ensembles in biomolecular recognition. *Nat. Chem. Biol.* 5: 789-796.

Bond, C.S., White, M.F., and Hunter, W.N. (2001). High resolution structure of the phosphohistidine-activated form of Escherichia coli cofactor-dependent phosphoglycerate mutase. *J. Biol. Chem.* 276: 3247-3253.

Bond, C.S., White, M.F., and Hunter, W.N. (2002). Mechanistic implications for Escherichia coli cofactor-dependent phosphoglycerate mutase based on the high-resolution crystal structure of a vanadate complex. *J. Mol. Biol.* 316: 1071-1081.

Botelho, H.M., Koch, M., Fritz, G., and Gomes, C.M. (2009). Metal ions modulate the folding and stability of the tumor suppressor protein S100A2. *FEBS J.* 276:1776-1786.

Boyer, P.D., Cross, R.L., and Momsen, W. (1973). A new concept for energy coupling in oxidative phosphorylation based on a molecular explanation of the oxygen exchange reactions. *PNAS* 70(10): 2837-2839.

Bradford, M.M. (1976). Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* 72: 248-254.

Bressi, J.C., Choe, J., Hough, M.T., Buckner, F.S., Van Voorhis, W.C., Verlinde, C.L.M.J., Hol, W.G.J., and Gelb, M.H. (2000). Adenosine analogues as inhibitors of Trypanosoma brucei phosphoglycerate kinase: Elucidation of a novel binding mode for a 2-Amino-N6-Substituted Adenosine. *J. Med. Chem.* 43: 4135-4150.

Brewer, J.M. (1981). Yeast enolase: Mechanism of activation by metal ions. *Crit. Rev. Biochem.*11(3): 209-254.

Brun, R., Blum, J., Chappuis, F., and Burri, C. (2010). Human African trypanosomiasis. *Lancet* 375: 148-159.

Bucher, D., Grant, B.J., Markwick, P.R., and McCammo, J.A. (2011). Accessing a hidden conformation of the maltose binding protein using accelerated molecular dynamics. *PLoS Comp. Biol.* 7(4): e1002034.

Bukrinsky, J. T., Bjerrum, M.J., and Kadziola, A. (1998). Native carboxypeptidase A in a new crystal environment reveals a different conformation of the important tyrosine 248. *Biochemistry* 37: 16555-16564.

Burgess, R. R. (2008). Protein purification, in Proteomics of the nervous system. Edited by Nothwang, H.G., and Pfeiffer, S.E. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. Accessed in February 2012. http://media.wiley.com/product_ data/excerpt/63/35273171/3527317163.pdf.

Bygrave, F.L. (1966). Studies on the interaction of metal ions with puruvate kinase from Ehrlich ascites-tumour cells and from rabbit muscle. *Biochem J.* 101(2): 488-494.

Callens, M., Kuntz, D.A., and Opperdoes, F.R. (1991). Kinetic properties of fructosebisphosphatealdolase from Trypanosoma brucei compared to aldolase from rabbit muscle and Staphylococcus aureus. *Mol. Biochem. Parasitol.* 47(1): 1-9.

Campbell, J.W., Watson, H.C., and Hodgson, G.I. (1974). Structure of yeast phosphoglycerate mutase. *Nature* 250: 301-303.

Carreras, J., Bartrons, R., and Grisolía, S. (1980). Vanadate inhibits 2,3-bisphosphoglycerate dependent phosphoglycerate mutases but does not affect the 2,3-bisphosphoglycerate independent phosphoglycerate mutases. *Biochem. Biophys. Res Commun.* 96(3): 1267-1273.

Castagnetto, J.M., Hennessy, S.W., Roberts, V.A., Getzoff, E.D., Tainer, J.A., and Pique, M.E. (2002). MDB: the Metalloprotein Database and Browser at the Scripps Research Institute. *Nucleic Acids Res.* 30: 379-382.

Castillo, E., Dea-Ayuela, M.A., Bolás-Fernández, F., Rangel, M., and González-Rosende, M.E. (2010). The kinetoplastid chemotherapy revisited: Current Drugs, Recent Advances and Future Perspectives. *Curr. Med. Chem.* 17: 4027-4051.

Cavaletto, M., Pessione, E., Vanni, A., and Giunta, C. (2000). Improved resistance to transition metals of a cobalt-substituted alcohol dehydrogenase 1 from Saccharomyces cerevisiae. *J. Biotech.* 84: 87-91.

Cavalli, A., Lizzi, F., Bongarzone, S., Belluti, F., Piazzi, L., and Bolognesi, M.L. (2010). Complementary medicinal chemistry-driven strategies toward new antitrypanosomal and antileishmanial lead drug candidates. *FEMS Immunol. Med. Microbiol.* 58(1): 51-60.

Chang, C., Chhor, G., Clancy, S., and Joachimiak, A. (2011). Crystal structure of phosphoglycerate mutase from Bacillus anthracis str. Sterne. PDB ID 3R7A. No publication yet available.

Chawla, B., and Madhubala, R. (2010). Drug targets in Leishmania. *J. Parasit. Dis.* 34(1): 1-13.

Chevalier, N., Rigden, D.J., Van Roy, J., Opperdoes, F.R., and Michels, P.A.M. (2000). Trypanosoma brucei contains a 2,3-bisphosphoglycerate independent phosphoglycerate mutase. *Eur. J. Biochem*. 267: 1464-1472.

Chiba, Y., Oshima, K., Arai, H., Ishii, M., and Igarashi, Y. (2012). Discovery and analysis of cofactor-dependent phosphoglycerate mutase homologs as novel phosphoserine phosphatases in Hydrogenobacter thermophilus. *J. Biol. Chem*. 287: 11934-11941.

Coffer, A.I., and Knowles, P.P. (1994). Divalent metal ions induce conformational change in pure, human wild-type p53 tumor suppressor protein. *Biochim. Biophys. Acta* 1209(2): 279-285.

Colasante, C., Robles, A., Li, C-H., Schwede, A., Benz, C., Voncken, F., Guilbride, D.L., and Clayton, C. (2007). Regulated expression of glycosomal phosphoglycerate kinase in Trypanosoma brucei. *Mol. Biochem. Parasitol.* 151(2): 193-204.

Coleman, J. E. (1965). Human carbonic anhydrase. Protein conformation and metal ion binding. *Biochemistry* 4(12): 2644-2655.

Coley, A.F., Dodson, H.C., Morris, M.T., and Morris, J.C. (2011). Glycolysis in the African trypanosome: Targeting enzymes and their subcellular compartments for therapeutic development. *Mol. Biol. Int.* 2011:123702.

Collet, J.-F., Stroobant, V., and Schaftingen, E.V. (2001). The 2,3-bisphosphoglycerate-independent phosphoglycerate mutase from Trypanosoma brucei: metal-ion dependency and phosphoenzyme formation. *FEMS Microbiol. Lett.* 204: 39-44.

Cowan, J. A. (1995). Introduction to the biological chemistry of magnesium. VCH Publishers, New York.

Cowan, J.A. (1997). Metal-mediated hydrolysis of biological phosphate ester hydrolysis. A critical analysis of the essential metal ion stoichiometry for nuclease activation. *J. Biol. Inorg. Chem.* 2: 168-176.

Cregg, J.M., Cereghino, J.L., Shi, J., and Higgins, D.R. (2000). Recombinant protein expression in Pichia pastoris. *Mol. Biotechnol.* 16(1): 23-52.

Cronin, C.N., and Tipton, K.F. (1987). Kinetic studies on the reaction catalysed by phosphofructokinase from Trypanosoma brucei. *Biochem. J.* 245: 13-18.

Crowhurst, G.S., Dalby, A.R., Isupov, M.N., Campbell, J.W., and Littlechild, J.A (1999). Structure of a phosphoglycerate mutase:3-phosphoglyceric acid complex at 1.7 Å. *Acta Crystallogr. D Biol. Crystallogr.* 55(11): 1822-1826.

Curdel, A., and Iwatsubo, M. (1968). Biosynthetic incorporation of cobalt into yeast alcohol dehydrogenase. *FEBS Lett.* 1: 133-136.

D'Alessio, G. (1999). The evolutionary transition from monomeric to oligomeric proteins: tools, the environment, hypotheses. *Prog. Biophys. Mol.Biol.* 72: 271-298.

Danel, F., Paetzel, M., Strynadka, N.C. and Page, M.G. (2001). Effect of divalent metal cations on the dimerization of OXA-10 and -14 class D beta-lactamases from Pseudomonas aeruginosa. Biochemistry 40(31):9412-9420.

Davies, D.R., Staker, B.L., Abendroth, J.A., Edwards, T.E., Hartley, R., Leonard, J., Kim, H., Rychel, A.L., Hewitt, S.N., Myler, P.J., and Stewart, L.J. (2011). An ensemble of structures of Burkholderia pseudomallei 2,3-bisphosphoglycerate-dependent phosphoglycerate mutase. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commu*. 67(9): 1044-1050.

Dax, C., Duffieux, F., Chabot, N., Coincon, M., Sygusch, J., Michels, P.A.M., and Blonski, C. (2006). Selective irreversible inhibition of fructose 1,6-bisphosphate aldolase from Trypanosoma brucei. *J. Med. Chem.* 49: 1499-1502.

de Walque, S., Opperdoes, F.R., and Michels, P.A.M. (1999). Cloning and characterization of Leishmania mexicana fructose-1,6-bisphosphate aldolase. *Mol. Biochem. Parasitol.* 103(2): 279-283.

DeLano, W. L. (2002). The PyMOL Molecular Graphics System, DeLano Scientific. San Carlos, CA, USA.

Djikeng, A., Raverdy, S., Foster, J., Bartholomeu, D., Zhang, Y., El-Sayed, N.M., and Carlow, C. (2007). Cofactor-independent phosphoglycerate mutase is an essential gene in procyclic form Trypanosoma brucei. *Parasitol. Res.* 100(4): 887-892.

Docampo, R. (1993) Calcium homeostasis in Trypanosoma cruzi. *Biol. Res.* 26: 189-196.

Docampo, R., de Souza, W., Miranda, K., Rohloff, P., and Moreno, S.N.J. (2005). Acidocalcisomes- Conserved from bacteria to man. *Nat. Rev. Microbiol.* 3: 251-261.

Dudev, T., and Lim, C. (2004). Monodentate versus bidentate carboxylate binding in magnesium and calcium proteins: What are the basic principles? *J. Phys. Chem. B* 108: 4546-4557.

Ellis, J. P., Culviner, P.H. and Cavagnero, S. (2009). Confined dynamics of a ribosome-bound nascent globin: Cone angle analysis of fluorescence depolarization decays in the presence of two local motions. *Protein Sci.* 18(10): 2003-2015.

Ernest, I., Callens, M., Uttaro, A.D., Chevalier, N., Opperdoes, F.R., Muirhead, H., and Michels, P.A.M. (1998). Pyruvate kinase of Trypanosoma brucei: overexpression, purification, and functional characterization of wild-type and mutated enzyme. *Protein Expr. Purif.* 13: 373-382.

Fenton, D. E. (1995). Biocoordination chemistry. Oxford University Press, Oxford.

Fielding, A. J., Kovaleva, E.G., Farquhar, E.R., Lipscomb, J.D., and Que, L. Jr. (2010). A hyperactive cobalt-substituted extradiol-cleaving catechol dioxygenase. *J. Biol. Inorg. Chem.* 16(2): 341-355.

Fieulaine, S., Boularot, A., Artaud, I., Desmadril, M., Dardel, F., Meinnel, T., and Giglione, C. (2011). Trapping conformational states along ligand-binding dynamics of peptide deformylase: The impact of induced fit on enzyme catalysis. *PLoS Biol.* 9(5): e1001066.

Finney, L.A., and O' Halloran. T.V. (2003). Transition metal speciation in the cell: Insights from the chemistry of metal ion receptors. *Science* 300: 931-936.

Fischer, E. (1894). Einfluss der Configuration auf die Wirkung der Enzyme. *Ber. Dtsch. Chem. Ges.* 27: 2985-2993.

Folta-Stogniew, E., and Williams, K.R. (1999). Determination of molecular masses of proteins in solution: Implementation of an HPLC size exclusion chromatography and laser light scattering service in a core laboratory. *J. Biomol. Tech.* 10(2): 51-63.

Foster, J.M., Raverdy, S., Ganatra, M.B., Colussi, P.A., Taron, C.H., and Carlow, C.K.S. (2009). The Wolbachia endosymbiont of Brugia malayi has an active phosphoglycerate mutase: a candidate target for anti-filarial therapies. *Parasitol. Res.*104(5): 1047-1052.

Fothergill-Gilmore, L.A., and Watson, H.C. (1989). The phosphoglycerate mutases. *Adv. Enzymol. Relat. Areas Mol. Biol.* 62: 227-313.

Fothergill-Gilmore, L.A., and Michels, P.A.M. (1993). Evolution of glycolysis. *Prog. Biophys Mol Biol.* 59(2): 105-235.

Fraser, H.I., Kratskhelia, M., and White, M.F. (1999). The two analogous phosphoglycerate mutases of Escherichia coli. *FEBS Lett.* 455: 344-348.

Fraústo da Silva, J.J.R., and Williams, R.J.P. (1991). The biological chemistry of the elements: The inorganic chemistry of life, 1st edition. Oxford University Press, Oxford.

Fraústo da Silva, J.J.R. and Williams, R.J.P. (2001). The biological chemistry of the elements: The inorganic chemistry of life, 2nd edition. Oxford University Press, Oxford.

Fuad, F.A.A., Fothergill-Gilmore, L.A., Nowicki, M.W., Eades, L.J., Morgan, H.P., McNae, I.W., Michels, P.A.M., and Walkinshaw, M.D. (2011). Phosphoglycerate mutase from Trypanosoma brucei is hyperactivated by cobalt in vitro, but not in vivo. *Metallomics* 3: 1310-1317.

Gaberc-Porekar, V., and Menart, V. (2001). Perspectives of immobilized-metal affinity chromatography. *J. Biochem. Biophys. Methods.* 49(1-3): 335-360.

Galperin, M.Y., Bairoch, A., and Koonin, E.V. (1998). A superfamily of metalloenzymes unifies phosphopentomutase and cofactor-independent phosphoglycerate mutase with alkaline phosphatases and sulfatases. *Protein Sci.* 7: 1829-1835.

Galperin, M.Y., and Jedrzejas, M.J. (2001). Conserved Core Structure and Active Site Residues in Alkaline Phosphatase Superfamily Enzymes. *Proteins* 45: 318-324.

Garcıá De La Torre, J., Huertas, M.L., and Carrasco, B. (2000). Calculation of hydrodynamic properties of globular proteins from their atomic-level structure. *Biophys. J.* 78: 719-730.

Geno Technology Inc., U. Cobalt Chelating Resin. Website accessed in February 2012. http://www.gbiosciences.com/ResearchProducts/Cobalt-Chelating-Resin.aspx.

Ghosh, S., Rasheedi, S., Rahim, S.S., Banerjee, S., Choudhary, R.K., Chakhaiyar, P., Ehtesham, N.Z., Mukhopadhyay, S., and Hasnain, S.E. (2004). Method for enhancing solubility of the expressed recombinant proteins in Escherichia coli. *Biotechniques* 37: 418-423.

Gilbert, I.H. (2002). Inhibitors of dihydrofolate reductase in leishmania and trypanosomes. *Biochim. Biophys. Acta* 1587: 249-257.

Gill, S. C. and von Hippel, P. H. (1989) Calculation of protein extinction coefficients from amino acid sequence data. *Anal. Biochem.* 182: 319-326.

Golgher, D., Vianna, C.H., and Moura, A.C. (2011). Drugs against leishmaniasis: Overview of market needs and pipeline. *Drug Dev. Res.* 72(6): 463-470.

Gomes, C.M., and Wittung-Stafshede, P. (2011). Protein Folding and Metal Ions. Mechanisms, Biology and Disease. CRC Press, Taylor & Francis Group, Boca Raton, FL.

Gomez-Ortiz, M., Gomis-Ruth, F.X., Huber, R., and Aviles, F.X. (1997). Inhibition of carboxypeptidase A by excess zinc: analysis of the structural determinant by X-ray crystallography. *FEBS Lett.* 400: 336-340.

Grana, X., Lecea, L., El-Maghrabi, M.R., Urena, J.M., Caellas, C., Carreras, J., Puigdomenech, P., Pilkis, S.J., and Climent, F. (1992). Cloning and sequencing of a cDNA encoding 2,3-bisphosphoglycerate-independent phosphoglycerate mutase from maize. *J. Biol. Chem.* 267: 12797-12803.

Gray, H. B. (2003). Biological inorganic chemistry at the beginning of the 21st century. *PNAS* 100(7): 3563-3568.

Gualdrón-López, M., Brennand, A., Hannaert, V., Quiñones, W., Cáceres, A.J., Bringaud, F., Concepción, J.L., and Michels, P.A.M. (2012). When, how and why glycolysis became compartmentalised in the Kinetoplastea. A new look at an ancient organelle. *Int. J. Parasitol.* 42(1): 1-20.

Guerra, D. G., Vertommen, D., Fothergill-Gilmore, L.A., Opperdoes, F.R., and Michels, P.A.M. (2004). Characterization of the cofactor-independent phosphoglycerate mutase from Leishmania mexicana mexicana. Histidines that coordinate the two metal ions in the active site show different susceptibilities to irreversible chemical modification. *Eur. J. Biochem.* 271(9): 1798-1810.

Gupta, S., Igoillo-Esteve, M., Michels, P.A.M., and Cordeiro, A.T. (2011). Glucose-6-phosphate dehydrogenase of trypanosomatids: Characterization, target validation, and drug discovery. *Mol. Biol. Int.* 2011: 135701.

Håkansson, K., and Wehnert, A. (1992). Structure of cobalt carbonic anhydrase complexed with bicarbonate. *J. Mol. Biol.* 228(4): 1212-1218.

Hall, D.R., Leonard, G.A., Reed, C.D., Watt, C.I., Berry, A., and Hunter, W.N. (1999). The crystal structure of Escherichia coli class II fructose-1,6-bisphosphate aldolase in complex with phosphoglycolohydroxamate reveals details of mechanism and specificity. *J. Mol. Biol.* 287(2): 383-394.

Hall, D.R., Bond, C.S., Leonard, G.A., Watt, C.I., Berry, A., and Hunter, W.N. (2002). Structure of tagatose-1,6-bisphosphate aldolase. Insight into chiral discrimination, mechanism, and specificity of class II aldolases. *J. Biol. Chem.* 277: 22018-22024.

Haraguchi, H. (2004). Metallomics as integrated biometal science. *J. Anal. At. Spectrom.*19: 5-14.

Harding, M.M., Nowicki, M.W., and Walkinshaw, M.D. (2010). Metals in protein structures: a review of their principal features. *Cryst. Rev.* 16(4): 247-302.

Hayward, S. (1999). Structural principles governing domain motions in proteins. *Proteins* 36: 425-435.

Hemdan, E.S., Zhao, Y.J., Sulkowski, E., and Porath, J. (1989). Surface topography of histidine residues: A facile probe by immobilized metal ion affinity chromatography. *PNAS* 86: 1811-1815.

Holden-Dye, L. and Walker, R.J. (2007). Anthelmintic drugs, *WormBook*, ed. The *C. elegans* Research Community, WormBook. doi/10.1895/wormbook.1.143.1. Website accessed in May 2012. http://www.wormbook.org.

Holz, R.C., Bzymek, K.P., and Swierczek, S.I. (2003). Co-catalytic metallopeptidases as pharmaceutical targets. *Curr. Opin. Chem. Biol.* 7(2): 197-206.

Hsin, K-Y., Sheng, Y., Harding, M.M., Taylor, P., and Walkinshaw, M.D. (2008). MESPEUS: a database of the geometry of metal sites in proteins. *J. Appl. Cryst.* 41: 963-968.

Hsin, K-Y., Morgan, H.P., Shave, S.R., Hinton, A.C.,Taylor, P., and Walkinshaw, M.D. (2011). EDULISS: a small-molecule database with data-mining and pharmacophore searching capabilities. *Nucleic Acids Res.* 39: D1042-D1048.

Huber, R., and Bennett., W.S.Jr. (1983). Functional significance of flexibility in proteins. *Biopolymers* 22(1): 261-279.

Huey, R., Morris, G.M., Olson, A.J., and Goodsell, D.S. (2007). A semiempirical free energy force field with charge-based desolvation. *J. Comp. Chem.* 28(6): 1145-1152.

Ibers, J.A., and Holm, R.H. (1980). Modeling coordination sites in metallobiomolecules. *Science* 209: 223-235.

Inglese, J., Auld, D.S., Jadhav, A., Johnson, R.L., Simeonov, A., Yasgar, A., Zheng, W., and Austin, C.P. (2006). Quantitative high-throughput screening: A titration-based approach that efficiently identifies biological activities in large chemical libraries. *PNAS* 103(31): 11473-11478.

Inglese, J., Shamu, C.E., and Guy, R.K. (2007a). Reporting data from high-throughput screening of small-molecule libraries. *Nat.Chem. Biol.* 3(8): 438-441.

Inglese, J., Johnson, R.L., Simeonov, A., Xia, M., Zheng, W., Austin, C.P., and Auld, D.S. (2007b). High-throughput screening assays for the identification of chemical probes. *Nat. Chem. Biol.* 3(8): 466-479.

Irving, H.M.N.H., and Williams, R.J.P. (1953). The stability of transition-metal complexes. *J. Chem. Soc*: 3192-3210.

Iyengar, G.V., Kollmer, W.E., and Bowen, H.J.M. (1978). The elemental composition of human tissues and body fluids. Verlag Chemie, Weinheim and New York.

Jacobs, R.T., Plattner, J.J., Nare, B., Wring, S.A., Chen, D., Freund, Y., Gaukel, E.G., Orr, M.D., Perales, J.B., Jenks, M., Noe, R.A., Sligar, J.M., Zhang, Y-K., Bacchi, C.J., Yarlett, N., and Don,R. (2011). Benzoxaboroles: a new class of potential drugs for human African trypanosomiasis. *Future Med. Chem.* 3(10): 1259-1278.

Jacobs, R.T., Nare, B.,Wring, S.S., Orr, M.D., Chen, D., Sligar, J.M., Jenks, M.X., Noe, R.A., Bowling, T.S., Mercer, L.T., Rewerts, C., Gaukel, E., Owens, J., Parham, R., Randolph, R., Beaudet, B., Bacchi, C.J.,Yarlett, N., Plattner, J.J., Freund, Y., Ding, C.,Akama, T., Zhang, Y-K., Brun, R., Kaiser, M., Scandale, I., and Don, R. (2011). SCYX-7158, an orally-active benzoxaborole for the treatment of stage 2 Human African Trypanosomiasis. *PLoS Negl. Trop. Dis*. 5(6): e1151.

Jedrzejas, M.J., Chander, M., Setlow, P., and Krishnasamy, G. (2000a). Structure and mechanism of action of a novel phosphoglycerate mutase from Bacillus stearothermophilus. *EMBO J.* 19(7): 1419-1431.

Jedrzejas, M.J., Chander, M., Setlow, P. and Krishnasamy, G. (2000b). Mechanism of catalysis of the cofactor-independent phosphoglycerate mutase from Bacillus stearothermophilus. Crystal structure of the complex with 2-phosphoglycerate. *J. Biol. Chem*. 275(30): 23146-23153.

Johnsen, U., and Schönheit, P. (2007). Characterization of cofactor-dependent and cofactor-independent phosphoglycerate mutases from Archaea. *Extremophiles* 11(5): 647-657.

Johnson, M., and Price, N.C. (1988). Do metal ions promote the re-activation of the 2,3-bisphosphoglycerate-independent phosphoglycerate mutases? *Biochem. J.* 252: 111-117.

Kar, L., Matsumura, P., and Johnson, M.E. (1992). Bivalent-metal binding to CheY protein. Effect on protein conformation. *Biochem. J.* 287: 521-531.

Karmi, O., Zayed, A., Baraghethi, S., Qadi, M. and Ghanem, R. (2011). Measurement of vitamin B12 concentration: A review on available methods. *The IIOAB J.* 2(2): 23-32.

Kachmar, J.F., and Boyer, P.D. (1953) Kinetic analysis of enzyme reactions. II. The potassium activation and calcium inhibition of pyruvic phosphoferase, *J. Biol. Chem.* 200: 669-683.

Kato, M., Chuang, J.L., Tso, S.C., Wynn, R.M., and Chuang, D.T. (2005). Crystal structure of pyruvate dehydrogenase kinase 3 bound to lipoyl domain 2 of human pyruvate dehydrogenase complex. *EMBO J. Bacteriol.* 24: 1763-1774.

Kempner, E.S. (1993). Movable lobes and flexible loops in proteins. Structural deformations that control biochemical activity. *FEBS Lett.* 326(1-3): 4-10.

Kim, E.E., and Wyckoff, H.W. (1991). Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis. *J. Mol. Biol.* 218(2): 449-464.

Klebe, G. (2009). Wirkstoffdesign: Entwurf und Wirkung von Arzneistoffen. Springer, Spektrum Academischer Verlag, Heidelberg. English translation is available online in http://pc1664.pharmazie.uni-marburg.de/resources/pdf/Chapter4.pdf .

Koshland, D.E. (1958). Application of a theory of enzyme specificity to protein synthesis. *PNAS* 44(2): 98-104.

Kost, T.A., Condreay, J.P., and Jarvis, D.L. (2005). Baculovirus as versatile vectors for protein expression in insect and mammalian cells. *Nat. Biotechnol*. 23(5): 567-575.

Kuhn, N.J., Setlow, B., Setlow, P., Cammack, R., and Williams, R. (1995). Cooperative manganese (II) activation of 3-phosphoglycerate mutase of Bacillus megaterium: a biological pH-sensing mechanism in bacterial spore formation and germination. *Arch. Biochem. Biophys.* 320: 35-42.

Kumar, A., Lomize, A., Jin, K.K., Carlton, D., Miller, M.D., Jaroszewski,L., Abdubek, P., Astakhova, T., Axelrod, H.L., Chiu, H-J., Clayton, T., Das, D., Deller, M.C., Duan, L., Feuerhelm, J., Grant, J.C., Grzechnik, A., Han, G.W., Klock, H.E., Knuth, M.W., Kozbial, P., Krishna, S.S., Marciano, D., McMullan, D., Morse, A.T., Nigoghossian, E., Okach, L., Reyes, R., Rife, C.L., Sefcovic, N.,Tien, H.J.,Trame, C.B., van den Bedem, H., Weekes, D., Xu, Q., Hodgson, K.O., Wooley, J., Elsliger, M-A., Deacon, A.M., Godzik, A., Lesley, S.A., and Wilsonad, I.A. (2009). Open and closed conformations of two SpoIIAA-like proteins (YP_749275.1 and YP_001095227.1) provide insights into membrane association and ligand binding. *Acta Crystallogr. Sect. F Struct. Biol .Cryst. Commun.* 66(10): 1245-1253.

Kumar, S., Ma, B., Tsai, C.J., Sinha, N., and Nussinov, R. (2000). Folding and binding cascades: dynamic landscapes and population shifts. *Protein Sci.* 9: 10-19.

Lackner, A., Genta, K., Koppensteiner, H., Herbacek, I., Holzmann, K., Spiegl-Kreinecker, S., Berger, W., and Grusch, M. (2008). A bicistronic baculovirus vector for transient and stable protein expression in mammalian cells. *Anal. Biochem.* 380(1): 146-148.

Lane, T.W., and Morel, F.M.M. (2000). Regulation of carbonic anhydrase expression by zinc, cobalt, and carbon dioxide in the marine diatom Thalassiosira weissflogii. *Plant Physiol.* 123(1): 345-352.

Larsen, T.M., Benning, M.M., Rayment, I., and Reed, G.H. (1998). Structure of the bis(Mg^{2+})-ATP-oxalate complex of the rabbit muscle pyruvate kinase at 2.1 A resolution: ATP binding over a barrel. *Biochemistry* 37: 6247-6255.

Lassila, J.K., and Herschlag, D. (2008). Promiscuous sulfatase activity and thio-effects in a phosphodiesterase of the alkaline phosphatase superfamily. *Biochemistry* 47(48): 12853-12859.

Leach, A.R., Shoichet, B.K., and Peishoff, C.E. (2006). Prediction of protein-ligand interactions. Docking and scoring: successes and gaps. *J. Med. Chem.* 49(20): 5851-5855.

Lee, B.H., and Nowak, T. (1992). Influence of pH on the Mn^{2+} activation of and binding to yeast enolase: A functional study. *Biochemistry* 31: 2165-2171.

Li, Z., Galvin, B.D., Raverdy, S., and Carlow, C.K.S. (2011). Identification and characterization of the cofactor-independent phosphoglycerate mutases of Dirofilaria immitis and its Wolbachia endosymbiont. *Vet. Parasitol.* 176(4): 350-356.

Li, Z., Xiong, F., Lin, Q., d'Anjou, M., Daugulis, A.J., Yang, D.S. and Hew, C.L. (2001). Low-temperature increases the yield of biologically active herring antifreeze protein in Pichia pastoris. *Protein Expr. Purif.* 21(3):438-445.

Lin, Z., Fernández-Robledo, J.S., Cellier, M.F.M., and Vasta, G.R. (2009). Metals and membrane metal transporters in biological systems: The role(s) of Nramp in host-parasite infections. *J. Argentine Chem. Soc.* 97(1): 210-225.

Lipinski, C.A., Lombardo, F., Dominy, B.W., and Feeney, P.J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Del. Rev.* 23(1-3): 3-25.

Lippard, S.J., and Berg, J.M. (1994). Principles of Bioinorganic Chemistry. University Science Books, Mill Valley, CA.

Liu, Y., and Eisenberg., D. (2002). 3D domain swapping: as domains continue to swap. *Protein Sci*. 11(6): 1285-1299.

López-Canut, V., Roca, M., Bertrán, J., Moliner, V., and Tuñón, I. (2011). Promiscuity in alkaline phosphatase superfamily. Unraveling evolution through molecular simulations. *J. Am. Chem. Soc.* 133(31): 12050-12062.

Makrides, S.V. (1996). Strategies for achieving high-level expression of genes in Escherichia coli. *Microbiol. Rev* 60:512-538.

Malvern Instruments. What is the hydrodynamic radius R_H? Website accessed in February 2012. http://www.imbb.forth.gr/people/aeconomou/pdf/hydrodynamic_radius.pdf.

Malys, N., Wishart, J.A., Oliver, S.G., and McCarthy, J.E. (2011). Protein production in Saccharomyces cerevisiae for systems biology studies. *Methods Enzymol.* 500: 197-212.

Maret, W. (2010). Metalloproteomics, metalloproteomes, and the annotation of metalloproteins. *Metallomics* 2: 117-125.

Marianayagam, N.J., Sunde, M., and Matthews, J.M. (2004). The power of two: protein dimerization in biology. *Trends Biochem. Sci.* 29: 618-625.

Marino, M.H. (1989). Expression systems for heterologous protein production. *BioPharm.* 2: 18-33.

Martinez-Oyanedel, J., McNae, I.W., Nowicki, M.W., Keillor, J.W., Michels, P.A.M., Fothergill-Gilmore, L.A., and Walkinshaw, M.D. (2007). The first crystal structure of phosphofructokinase from a eukaryote: Trypanosoma brucei. *J. Mol. Biol.* 366(4): 1185-1198.

Mayence, A., Vanden Eynde, J.J., LeCour L, Jr., Walker, L.A., Tekwani, B.L. and Huang, T.L. (2004). Piperazine-linked bisbenzamidines: a novel class of antileishmanial agents. *Eur. J. Med. Chem.* 39(6):547-553.

McMillen, L., Beacham, I.R., and Burns, D.M. (2003). Cobalt activation of Escherichia coli 5´-nucleotidase is due to zinc ion displacement at only one of two metal-ion-binding sites. *Biochem. J.* 372: 625-630.

McNae, I.W., Martinez-Oyanedel, J., Keillor, J.W., Michels, P.A.M., Fothergill-Gilmore, L.A., and Walkinshaw, M.D. (2009). The crystal structure of ATP-bound phosphofructokinase from Trypanosoma brucei reveals conformational transitions different from those of other phosphofructokinases. *J. Mol. Biol.* 385(5): 1519-1533.

Mehio, W., Kemp, G.J.L., Taylor, P., and Walkinshaw, M.D. (2010). Identification of protein binding surfaces using surface triplet propensities. *Bioinformatics* 26 (20): 2549-2555.

Mercaldi, G., Pereira, H., Cordeiro, A., Michels, P.A.M., and Thiemann, O. (2012). Phophoglycerate mutase from Trypanosoma brucei: structure and catalytic mechanism. *FEBS J.* 279(11): 2012-2021.

Michels, P.A.M., Chevalier, N., Opperdoes, F.R., Rider, M.H., and Rigden, D.J. (1997). The glycosomal ATP-dependent phosphofructokinase of Trypanosoma brucei must have evolved from an ancestral pyrophosphate-dependent enzyme. *Eur. J. Biochem.* 250(3): 698-704.

Mildvan, A.S., and Cohn, M. (1965). Kinetic and magnetic resonance studies of the pyruvate kinase reaction. I. Divalent metal complexes of pyruvate kinase. *J. Biol. Chem*. 240: 238-246.

Minodier, P., and Parola, P. (2007). Cutaneous leishmaniasis treatment. *Travel Med. Infect. Dis.* 5(3): 150-158.

Misset, O., and Opperdoes, F.R. (1984). Simultaneous purification of hexokinase, class-I fructose-bisphosphate aldolase, triosephosphate isomerase and phosphoglycerate kinase from Trypanosoma brucei. *Eur. J. Biochem*. 144: 475-483.

Morgan, H.P., McNae, I.W., Nowicki, M.W., Zhong, W., Michels, P.A.M, Auld, D.S., Fothergill-Gilmore, L.A., and Walkinshaw, M.D. (2011). The trypanocidal drug suramin and other trypan blue mimetics are inhibitors of pyruvate kinases and bind to the adenosine site. *J. Biol. Chem.* 286(36): 31232-31240.

Moriguchi, I., Hirono, S., Liu, Q., Nakagome, Y., and Matsushita, Y. (1992). Simple methods of calculating octanol water partition coefficient. *Chem. Pharm. Bull.* 40: 127-130.

Morris, G.M., Goodsell, D. S., Halliday, R.S., Huey, R., Hart, W.E., Belew, R.K., and Olson, A.J.(1998). Automated docking using a Lamarckian Genetic Algorithm and empirical binding free energy function. *J. Comput. Chem.* 19: 1639-1662.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S. and Olson, A.J. (2009). AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* 30: 2785-2791.

Müller, P., Sawaya, M.R., Pashkov, I., Chan, S., Nguyen, C., Wu, Y., Perry, L.J., and Eisenberg, D. (2005). The 1.70 angstroms X-ray crystal structure of Mycobacterium tuberculosis phosphoglycerate mutase. *Acta Crystallogr. D Biol.Crystallogr.* 61(3): 309-315.

Murray, H. W., Berman, J.D., Davies, C.R., and Saravia, N.G. (2005). Advances in leishmaniasis. *Lancet* 366: 1561-1577.

Nallamsetty, S., and Waugh, D.S. (2007). Mutations that alter the equilibrium between open and closed conformations of Escherichia coli maltose-binding protein impede its ability to enhance the solubility of passenger proteins. *Biochem. Biophys. Res. Commun.* 364(3): 639-644.

Nar, H., Huber, R., Messerschmidt, A., Filippou, A.C., Barth, M., Jaquinod, M., van de Kamp, M., and Canters. G.W. (1992). Characterization and crystal structure of zinc azurin, a by-product of heterologous expression in Escherichia coli of Pseudomonas aeruginosa copper azurin. *Eur. J. Biochem.* 205: 1123-1129.

National Diagnostics. Native Protein Electrophoresis. Website accessed in February 2012. http://www.nationaldiagnostics.com/article_info.php/articles_id/60.

Nowicki, M.W., Kuaprasert, B., McNae, I.W., Morgan, H.P., Harding, M.M., Michels, P.A.M., Fothergill-Gilmore, L.A., and Walkinshaw, M.D. (2009). Crystal structures of Leishmania mexicana phosphoglycerate mutase suggest a one-metal mechanism and a new enzyme subclass. *J. Mol. Biol.* 394: 535-543.

Nukui, M., Mello, L.V., Littlejohn, J.E., Setlow, B., Setlow, P., Kim, K., Leighton, T., and Jedrzejas, M.J. (2007). Structure and molecular mechanism of Bacillus anthracis cofactor-independent phosphoglycerate mutase: A crucial enzyme for spores and growing cells of Bacillus species. *Biophy. J.* 92(3): 977-988.

Oduro, K.K., Flynn, I.W., and Bowman, I.B.R. (1980). Trypanosoma brucei: activities and subcellular disruption of glycolytic enzymes from differentially disrupted cells. *Exp. Parasitol.* 50:123-135.

Olsson, U., and Wolf-Watz, M. (2010). Overlap between folding and functional energy landscapes for adenylate kinase conformational change. *Nat. Commun.* 1(8): 111.

Opperdoes, F.R., and Borst, P. (1977). Localization of nine glycolytic enzymes in a microbody-like organelle in Trypanosoma brucei: The glycosome. *FEBS Lett.* 80: 360-364.

Ortega, A., Amorós, D., and Garciá de la Torre, J. (2011). Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. *Biophys. J.* 101: 892-898.

Ortega, A., and Garciá de la Torre, J. (2007). Equivalent radii and ratios of radii from solution properties as indicators of macromolecular conformation, shape, and flexibility. *Biomacromolecules* 8: 2464-2475.

Outten, C.E., and O'Halloran, T.V (2001). Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. *Science* 292(5526): 2488-2492

Owino, A.V., Masiga, K.D., and Limo, K.M. (2008). RNA interference: a pathway to drug target identification and validation in trypanosome. *African J. Biochem. Res.* 2(3): 66-73.

Page, M.J., and Di Cera, E. (2006). Role of Na^+ and K^+ in enzyme function. *Physiol. Rev.* 86: 1049-1092.

Parisi, A.F., and Vallee, B.L. (1969). Zinc metalloenzymes: Characteristics and significance in biology and medicine. *American J. Clin. Nutr.* 22(9): 1222-1230.

Park, C., and Raines, R.T. (2000). Dimer formation by a "monomeric" protein. *Protein Sci.* 9: 2026-2033.

Permyakov, E.A., and Kretsinger, R.H. (2009). Cell signaling, beyond cytosolic calcium in eukaryotes. *J. Inorg. Biochem.* 103(1): 77-86.

Perola, E., Walters, W.P., and Charifson, P.S. (2004). A detailed comparison of current docking and scoring methods on systems of pharmaceutical relevance. *Proteins* 56: 235-249.

Petsko, G. A., and Ringe, D. (2004). Protein structure and function, in Primers in Biology. New Science Press, Mill Valley, C.A.

Poonperm, B., Guerra, D.G., McNae, I.W., Fothergill-Gilmore, L.A., and Walkinshaw, M.D. (2003). Expression, purification, crystallization and preliminary crystallographic analysis of Leishmania mexicana phosphoglycerate mutase. *Acta Crystallogr. D Biol. Crystallogr.* 59(7): 1313-1316.

Poonperm, B. (2005). X-ray crystallographic and ligand binding studies of the drug target phosphoglycerate mutase from Leishmania mexicana and Trypanosoma brucei. PhD thesis. Institute of Structural and Molecular Biology, School of Biological Sciences. University of Edinburgh, Edinburgh.

Porath, J. (1992). Immobilized metal affinity chromatography. *Protein Expr. Purif.* 3: 263-281.

Pozdnyakova, I., Guidry, J., and Wittung-Stafshede, P (2000). Copper triggered b-hairpin formation: initiation site for azurin folding? *J. Am.Chem. Soc.* 122: 6337-6338.

Prescott, J.M., Wagner, F.W., Holmquist, B., and Vallee, B.L. (1985). Spectral and kinetic studies of metal-substituted Aeromonas aminopeptidase: nonidentical, interacting metal-binding sites. *Biochemistry* 24(20): 5350-5356.

Price, N.C., and Nairn, J. (2009). Exploring proteins: a student's guide to experimental skills and methods. Oxford University Press Inc., New York.

Promega. Kinase-Glo® Luminescent Kinase Assays. Website accessed in March 2012. http://www.promega.com/products/cell-signaling/protein--kinases-and-kinase-assays/kinase_glo-luminescent-kinase-assays/.

Quiñones, W., Peñaa, P., Domingo-Sananesa, M., Cáceresa, A., Michels, P.A.M., Avilana, L., and Concepció, J.L. (2007). Leishmania mexicana: Molecular cloning and characterization of enolase. *Exp. Parasitol.* 116(3): 241-251.

Ramakrishnan, B., Boeggeman, E., and Qasba, P.K. (2004). Effect of the Met344His mutation on the conformational dynamics of bovine beta-1,4-galactosyltransferase: crystal structure of the Met344His mutant in complex with chitobiose. *Biochemistry* 43: 12513-12522

Raverdy, S., Zhang, Y., Foster, J., and Carlow, C.K. (2007). Molecular and biochemical characterization of nematode cofactor independent phosphoglycerate mutases. *Mol. Biochem. Parasitol.* 156(2): 210-216.

Rej, R., and Bretaudiere, J-P. (1980). Effects of metal ions on the measurement of alkaline phosphatase activity. *Clin. Chem.* 26(3): 423-428

Remington, S., Wiegand, G. and Huber, R. (1982). Crystallographic refinement and atomic models of two different forms of citrate synthase at 2.7 and 1.7 A resolution. *J. Mol Biol.* 158(1): 111-152.

Rester, U. (2008). From virtuality to reality - Virtual screening in lead discovery and lead optimization: A medicinal chemistry perspective. *Curr. Opin. Drug Discov. Devel.* 11(4): 559-568.

Rigden, D.J., Phillips, S.E.V., Michels, P.A.M., and Fothergill-Gilmore, L.A. (1999). The structure of pyruvate kinase from Leishmania mexicana reveals details of the allosteric transition and unusual effector specificity. *J. Mol. Biol.* 291(3): 615-635.

Rigden, D.J., Bagyan, I., Lamani, E., Setlow, P., and Jedrzejas, M.J. (2001). A cofactor-dependent phosphoglycerate mutase homolog from Bacillus stearothermophilus is actually a broad specificity phosphatase. *Protein Sci.* 10(9): 1835-1846.

Rigden, D.J. (2008). The histidine phosphatase superfamily: structure and function. *Biochem J.* 409(2): 333-348.

Robinson, C.R., Liu, Y., Thomson, J.A., Sturtevant, J.M., and Sligar, S.G. (1997). Energetics of heme binding to native and denatured states of cytochrome b562. *Biochemistry* 36(51): 16141-16146.

Rollinger, J.M., Stuppner, H., and Langer, T. (2008). Virtual screening for the discovery of bioactive natural products. *Prog. Drug Res.* 65(211): 213-249.

Romano, A.H., and Conway, T. (1996). Evolution of carbohydrate metabolic pathways. *Res Microbiol.* 147(6-7): 448-455.

Rosser, M.P., Xia, W., Hartsell, S., McCaman, M., Zhu, Y., Wang, S., Harvey, S., Bringmann, P., and Cobb, R.R. (2005). Transient transfection of CHO-K1-S using serum-free medium in suspension: a rapid mammalian protein expression system. *Protein Expr. Purif.* 40(2): 237-243.

Rulísek, L., and Vondrásek, J. (1998). Coordination geometries of selected transition metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}, Cd^{2+}, and Hg^{2+}) in metalloproteins. *J. Inorg. Biochem.* 71(3-4): 115-127.

Ruzheinikov, S.N., Das, S.K., Sedelnikova, S.E., Baker, P.J., Artymiuk, P.J., Garcia-Lara, J., Foster, S.J., and Rice, D.W. (2004). Analysis of the open and closed conformations of the GTP-binding protein YsxC from Bacillus subtilis. *J. Mol. Biol.* 339: 265-278.

Scott, D.A., Docampo, R., Dvorak, J.A., Shi, S., and Leapman, R.D. (1997). In situ compositional analysis of acidocalcisomes in Trypanosoma cruzi. *J. Biol. Chem.* 272(44): 28020-28029.

Seed, J.R., and Baquero, M.A. (1965). The characterization of hexokinase from Trypanosoma rhodesiense and Trypanosoma gambiense. *J. Eukaryotic Microbiol.* 12(3): 427-432.

Sharlow, E.R., Lyda, T.A., Dodson, H.C., Mustata, G., Morris, M.T., Leimgruber, S.S., Lee, K-H., Kashiwada, Y., Close, D., Lazo, J.S., and Morris, J.C. (2010). A target-based high throughput screen yields Trypanosoma brucei hexokinase small molecule inhibitors with antiparasitic activity. *PLoS Negl. Trop. Dis.* 4(4): e659.

Shave, S. R. (2010). Development of high performance structure and ligand based virtual screening techniques. PhD thesis. Institute of Structural and Molecular Biology, School of Biological Sciences. University of Edinburgh, Edinburgh.

Sherman, F., Stewart, J.W., and Tsunasawa, S. (1985) Methionine or not methionine at the beginning of a protein. *BioEssays* 3: 27-31.

Shim, J., and Mackerell., A.D.Jr. (2011). Computational ligand-based rational design: Role of conformational sampling and force fields in model development. *Medchemcomm.* 2(5): 356-370.

Silverman, J.S., and Bangs, J.D. (2012). Form and function in the trypanosomal secretory pathway. *Curr. Opin. Microbiol.* In press.

Singh, S. (2006). New developments in diagnosis of leishmaniasis. *Indian J. Med. Res.* 123: 311-330.

Smith, G.C., McWilliams, A.D., and Hass, L.F. (1986). Wheat germ phosphoglycerate mutase: evidence for a metalloenzyme. *Biochem. Biophys. Res. Commun.* 136(1): 336-340.

Sommer, J.M., and Wang, C.C. (1994). Targeting proteins to the glycosomes of African trypanosomes. *Annu. Rev. Microbiol.*(48): 105-138.

Sparkman, O. D. (2000). Mass spectrometry desk reference. Global View Pub, Pittsburgh.

Spector, S., Flynn, J.M., Tidor, B., Baker, T.A., and Sauer, R.T. (2003). Expression of N-formylated proteins in Escherichia coli. *Protein Expr. Purif.* 32: 317-322.

Steverding, D. (2008). The history of African trypanosomiasis. *Parasit. Vectors* 1:3.

Sugiura, T., and Noguchi, Y. (2009). Substrate-dependent metal preference of PPM1H, a cancer-associated protein phosphatase 2C: comparison with other family members. *BioMetals* 22(3): 469-477.

Sundar, S., and Chatterjee., M. (2006). Visceral leishmaniasis - current therapeutic modalities. *Indian J. Med. Res.* 123:345-352.

Tainer, J.A., Getzoff, E.D., Beem, K.M., Richardson, J.S., and Richardson, D.C. (1982). Determination and analysis of the 2 A-structure of copper, zinc superoxide dismutase. *J. Mol Biol.* 160(2): 181-217.

Tainer, J.A., Roberts, V.A., and Getzoff, E.D. (1992). Protein metal-binding sites. *Curr. Opin. Biotechnol.* 3: 378-387.

Tam, M., Gómez, S., Gonzalez-Gross, M., and Marcos, A. (2003). Possible roles of magnesium on the immune system. *Eur. J. Clin. Nutr.* 57: 1193-1197.

Tama, F., and Sanejouand, Y.H. (2001). Conformational change of proteins arising from normal mode calculations. *Protein Eng.* 14(1): 1-6.

Taylor, P., Blackburn, E., Sheng, Y.G., Harding, S., Hsin, K-Y., Kan, D., Shave, S., and Walkinshaw, M.D. (2008). Ligand discovery and virtual screening using the program LIDAEUS. *British J. Pharm.* 153(S1): S55-S67.

Temperini, C., Scozzafava, A., and Supuran, C.T. (2008). Carbonic anhydrase activation and the drug design. *Curr. Pharm. Des.* 14(7): 708-715.

Thompson, L.C., Goswami, S., Ginsberg, D.S., Day, D.E., Verhamme, I.M., and Peterson, C.B. (2011). Metals affect the structure and activity of human plasminogen activator inhibitor-1. I. Modulation of stability and protease inhibition. *Protein Sci.* 20: 353-365.

Tottey, S., Waldron, K.J., Firbank, S.J., Reale, B., Bessant, C., Sato, K., Cheek, T.R., Gray, J., Banfield, M.J., Dennison, C., and Robinson, N.J. (2008). Protein folding location can regulate manganese-binding versus copper- or zinc-binding. *Nature* 455(7216): 1138-1142.

Trott, O., and Olson, A.J. (2010). AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* 31: 455-461.

Trunz, B.B., Jędrysiak, R.,Tweats, D., Brun, R., Kaiser, M., Suwiński, J., and Torreele, E. (2011). 1-Aryl-4-nitro-1H-imidazoles, a new promising series for the treatment of human African trypanosomiasis. *Eur. J. Med. Chem.* 46(5): 1524-1535.

Valberg, L.S., Holt, J.M., Paulson, E., and Szivek, J. (1965). Spectrochemical analysis of sodium, potassium, calcium, magnesium, copper, and zinc in normal human erythrocytes. *J. Clin. Invest.* 44(3): 379-389.

Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D.P., Lins, L., Van Den Abbeele, J., Pays, A., Tebabi, P., Van Xong, H., Jacquet, A., Moguilevsky, N., Dieu, M., Kane, J.P., De Baetselier, P., Brasseur, R., and Pays, E. (2003). Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature* 422: 83-87.

Verlinde, C.L.M.J., Hannaert, V., Blonski, C., Willson, M., Périé, J.J., Fothergill-Gilmore, L.A., Opperdoes, F.R., Gelb, M.H., Hola, W.G.J., and Michels, P.A.M. (2001). Glycolysis as a target for the design of new anti-trypanosome drugs. *Drug Resist. Updat.* 4(1): 50-65.

Vetter, I.R., and Wittinghofer, A. (1999). Nucleoside triphosphate-binding proteins: different scaffolds to achieve phosphoryl transfer. *Q. Rev. Biophys.* 32: 1-56.

Villanueva, J., Hoshino, M., Katou, H., Kardos, J., Hasegawa, K., Naiki, H., and Goto, Y. (2004). Increase in the conformational flexibility of beta 2-microglobulin upon copper binding: a possible role for copper in dialysis-related amyloidosis. *Protein Sci.* 13(3): 797-809.

von Geldern, T., Harhay, M.O., Scandale, I., and Don, R. (2011). Kinetoplastid parasites. *Top. Med. Chem.* 7: 181-241.

Wang, J-L., Walling, L.L., Jauh, G.Y., Gu, Y-Q., and Lord, E.M. (1996). Lily cofactor-independent phosphoglycerate mutase: purification, partial sequencing, and immunolocalization. *Planta* 200(3): 343-352.

Wang, Y., Wei, Z., Liu, L., Cheng, Z., Lin, Y., Ji, F., and Gong, W. (2005). Crystal structure of human B-type phosphoglycerate mutase bound with citrate. *Biochem. Biophys. Res. Commun.* 331(4): 1207-1215.

Wang, Y., Lam, K.S.L., Yau, M., and Xu, A. (2008). Post-translational modifications of adiponectin: mechanisms and functional implications. *Biochem. J.* 409: 623-633.

Watabe K., and Freese, E. (1979). Purification and properties of the manganese dependent phosphoglycerate mutase of Bacillus subtilis. *J Bacteriol.* 137: 773-778.

WHO. Human African trypanosomiasis (sleeping sickness). Website accessed in March 2012. http://www.who.int/mediacentre/factsheets/fs259/en/.

WHO. Leishmaniasis - Burden of disease. Website accessed in March 2012. http://www.who.int/leishmaniasis/ burden/en/.

Wiegand, G., and Remington, S.J. (1986). Citrate synthase: structure, control, and mechanism. *Annu. Rev. Biophys. Biophys. Chem.* 15: 97-117.

Wiese, M., Morris, A., and Grant, K.M. (2009). Trypanosomatid protein kinases as potential drug targets. In: Antiparasitic and antibacterial drug discovery: from molecular targets to drug candidates. Drug Discovery in Infectious Diseases, Wiley-VCH, Weinheim: 227-247.

Wittung-Stafshede, P. (2002). Role of cofactors in protein folding. *Acc. Chem. Res.* 35(4): 201-208.

Wu, N., and Veillette., A. (2011). Immunology: Magnesium in a signalling role. *Nature* 475(7357): 471-476.

Wu, S., Bottini, M., Rickert, R.C., Mustelin, T., and Tautz, L. (2009). In silico screening for PTPN22 inhibitors: Active hits from an inactive phosphatase conformation. *Chem. Med. Chem.* 4: 440-444.

Xiong, Z-H., and Ruben, L. (1998). Trypanosoma brucei: The dynamics of calcium movement between the cytosol, nucleus, and mitochondrion of intact cells. *Exp. Parasitol.* 88(3): 231-239.

Yee, D., and Morel, F.M.M. (1996). In vivo substitution of zinc by cobalt in carbonic anhydrase of a marine diatom. *Limnol. Oceanogr.* 41: 573-577.

Zhang, Y., Dougherty, M., Downs, D.M., and Ealick, S.E. (2004). Crystal structure of an aminoimidazole riboside kinase from Salmonella enterica: implications for the evolution of the ribokinase superfamily. *Structure* 12: 1809-1821.

Zhao, Z., and Assman, S.M. (2011). The glycolytic enzyme, phosphoglycerate mutase, has critical roles in stomatal movement, vegetative growth, and pollen production in Arabidopsis thaliana. *J. Exp. Botany.* 62 (14): 5179-5189.

Zijlstra, E.E., Musa, A.M., Khalil, E.A.G., El Hassan, I.M., and El- Hassan, A.M. (2003). Post -kala-azar dermal leishmaniasis. *Lancet Infect. Dis.* 3: 87-98.

Ziv, G., Thirumalaib, D., and Haran, G. (2008). Collapse transition in proteins. *Phys. Chem. Chem. Phys.* 11: 83-93.

Publication

Fuad, F.A.A., Fothergill-Gilmore, L.A., Nowicki, M.W., Eades, L.J., Morgan, H.P., McNae, I.W., Michels, P.A.M., and Walkinshaw, M.D. (2011). Phosphoglycerate mutase from Trypanosoma brucei is hyperactivated by cobalt in vitro, but not in vivo. *Metallomics* 3: 1310-1317.

Metallomics

Cite this: *Metallomics*, 2011, **3**, 1310–1317

www.rsc.org/metallomics

PAPER

Phosphoglycerate mutase from *Trypanosoma brucei* is hyperactivated by cobalt *in vitro*, but not *in vivo*

Fazia Adyani Ahmad Fuad,[a] Linda A. Fothergill-Gilmore,[a] Matthew W. Nowicki,[a] Lorna J. Eades,[b] Hugh P. Morgan,[a] Iain W. McNae,[a] Paul A. M. Michels[c] and Malcolm D. Walkinshaw*[a]

Received 28th July 2011, Accepted 22nd September 2011
DOI: 10.1039/c1mt00119a

Production of ATP by the glycolytic pathway in the mammalian pathogenic stage of protists from the genus *Trypanosoma* is required for the survival of the parasites. Cofactor-independent phosphoglycerate mutase (iPGAM) is particularly attractive as a drug target because it shows no similarity to the corresponding enzyme in humans, and has also been genetically validated as a target by RNAi experiments. It has previously been shown that trypanosomatid iPGAMs require Co^{2+} to reach maximal activity, but the biologically relevant metal has remained unclear. In this paper the metal content in the cytosol of procyclic and bloodstream-form *T. brucei* (analysed by inductively coupled plasma-optical emission spectroscopy) shows that Mg^{2+}, Zn^{2+} and Fe^{2+} were the most abundant, whereas Co^{2+} was below the limit of detection (<0.035 μM). The low concentration indicates that Co^{2+} is unlikely to be the biologically relevant metal, but that instead, Mg^{2+} and/or Zn^{2+} may assume this role. Results from metal analysis of purified *Leishmania mexicana* iPGAM by inductively coupled plasma-mass spectrometry also show high concentrations of Mg^{2+} and Zn^{2+}, and are consistent with this proposal. Our data suggest that *in vivo* cellular conditions lacking Co^{2+} are unable to support the maximal activity of iPGAM, but instead maintain its activity at a relatively low level by using Mg^{2+} and/or Zn^{2+}. The physiological significance of these observations is being pursued by structural, biochemical and biophysical studies.

Introduction

Metals are abundant in biological macromolecules, and about one-third of all proteins contain metals that are essential for diverse functions such as enzyme activity, gene expression and cell signalling, as well as for the maintenance of structural integrity. The burgeoning discipline of metalloproteomics is providing insights into the wide range of functions of metalloproteins, and into the common structural signatures encountered in protein metal-binding sites.[1–3] These studies are also sounding notes of caution concerning the possibilities for mis-identification of the native metal within a particular metalloprotein because of (i) lack of selectivity of metal-binding sites, (ii) variations in metal concentrations in different environments and (iii) artefactual problems encountered during protein purification. Replacement of a native metal by an alternative metal can have a substantial impact on the properties of the metalloprotein, and in the cases of metalloenzymes may even strikingly enhance the activity above that of the native enzyme. The replacement of Zn^{2+} by Co^{2+}, Ni^{2+} or Cu^{2+} in aminopeptidase is an example in which the non-native metals enhance activity by up to 10-fold.[4] Developments in atomic spectroscopy have enabled trace amounts of metals to be detected in scarce biological and purified protein samples,[5,6] and therefore have encouraged studies to identify the biologically relevant native metal. Inductively coupled plasma-optical emission spectroscopy (ICP-OES) and inductively coupled plasma-mass spectrometry (ICP-MS) have been used in this study to measure the metal concentrations in subcellular samples derived from protist parasites and in purified protein samples, respectively.

The glycolytic pathway of trypanosomatid parasites has been identified as a possible drug target because the production of ATP in the pathogenic bloodstream stage of trypanosomes is entirely dependent on glycolysis.[7] In particular, the glycolytic enzyme cofactor-independent phosphoglycerate mutase (iPGAM) is an attractive target because it is completely different from the cofactor-dependent enzyme that catalyses the corresponding reaction in humans,[8] and because it has also been genetically

[a] *Structural Biochemistry Group, Institute of Structural and Molecular Biology, University of Edinburgh, Michael Swann Building, The King's Buildings, Mayfield Road, Edinburgh, UK EH9 3JR*
[b] *School of Chemistry, University of Edinburgh, Joseph Black Building, The King's Buildings, West Mains Road, Edinburgh, UK EH9 3JJ*
[c] *Research Unit for Tropical Diseases, de Duve Institute and Laboratory of Biochemistry, Université catholique de Louvain, Avenue Hippocrate 74, B-1200 Brussels, Belgium*

validated as a drug target.[9] Phosphoglycerate mutase (PGAM) catalyses the interconversion between 3-phosphoglycerate and 2-phosphoglycerate in the glycolytic and gluconeogenic pathways. Cofactor-independent PGAM (**i**PGAM) is active in the absence of the cofactor 2,3-bisphosphoglycerate, and is found in protists including *Trypanosoma brucei* and *Leishmania mexicana*, as well as in plants and in some nematodes, fungi, algae, bacteria and archaea. All experimentally characterised iPGAMs are monomeric with a mass of about 60 000 Da, and have been classified as members of the metal-dependent alkaline phosphatase superfamily.[10] By contrast, cofactor-dependent PGAM (**d**PGAM) found in humans and other vertebrates, and in some yeasts and bacteria requires the bisphosphoglycerate cofactor for activity (but no metals) and is usually dimeric or tetrameric with subunits of about 27 000 Da.[11] The two types of PGAM are unrelated and share no sequence similarity.

There is an absolute requirement for divalent metals by iPGAM, and the enzyme is inactive after treatment with metal chelators.[12–14] The crystal structures of iPGAMs from *Bacillus stearothermophilus* (*Bs* iPGAM)[15] and from *L. mexicana* (*Lm* iPGAM)[16] share 33% sequence identity, and in their closed conformations possess essentially identical topologies with all 16 active-site residues being the same (Fig. 1). The additional availability of the crystal structure of iPGAM from *Bacillus anthracis* (*Ba* iPGAM, 78% sequence identity to *Bs* iPGAM) in a strikingly different open conformation has enabled detailed structural and molecular dynamic comparisons between the two *Bacillus* iPGAMs.[17] The two metal binding sites illustrated in Fig. 1 have different coordination geometries in these open and closed iPGAM structures, indicating the essential roles played by the metals not only for catalysis, but also for enzyme conformation. Interestingly, the specific metal required by iPGAMs varies from one organism to another, despite overall conservation of the structure. *Bs* and *Ba* iPGAMs require Mn^{2+},[18] but *T. brucei* iPGAM (*Tb* iPGAM) and *Lm* iPGAM prefer Co^{2+},[19,20] although other divalent metals may support low activity of trypanosomatid iPGAMs. In addition, Mg^{2+} is favoured by nematode iPGAMs,[21] and Zn^{2+} is required by alkaline phosphatase, a more-distant member of the metal-dependent phosphatase family.[22] Plant iPGAMs are also known to depend on divalent metals, but the identities of the physiologically relevant metals remain controversial.

Co^{2+} is rare in nature, and it has been suggested that it, like Ni^{2+}, may in fact, represent a relic of early life before the advent of dioxygen.[23] In the reducing and sulphide-rich conditions of early life, the availability of Zn^{2+} was low because of its precipitation as a sulphide, but Co^{2+} and Ni^{2+} could have acted as valuable metal biocatalysts. The change to a dioxygen environment encouraged a switch to metal catalysts such as Mn^{2+} and Zn^{2+}. In these circumstances, it is possible to envisage that an enzyme active site that had originally evolved to use Co^{2+} may also be able to bind Zn^{2+}, but would function with lower catalytic efficiency. It is relevant to note in this context that epr, nmr and X-ray crystallographic studies of metal binding sites such as those in carbonic anhydrase have shown that it is particularly easy to substitute Zn^{2+} by other metals such as Co^{2+}.[24–27] In these circumstances it is likely that the 'native' metal used by a particular enzyme may be more a function of the concentration of acceptable divalent metals in its surroundings than of the selectivity of the enzyme metal binding site.

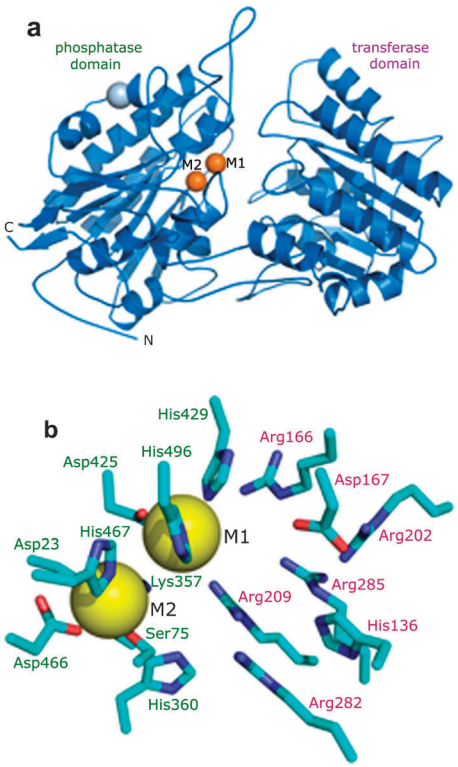


Fig. 1 Structure of iPGAM as illustrated by the enzyme from *L. mexicana*.[16] (a) A monomer of iPGAM is folded into two distinct domains designated phosphatase and transferase. The two divalent metal ions at the active site are shown by gold spheres. A Na^+ ion (grey sphere) is distant from the active site, and probably plays a structural role. (b) The active site of iPGAM crystallised in the presence of 4 mM Co^{2+}. The 16 amino acid residues shown (Asp23 and Lys357 are mostly hidden by the sphere of M2) are highly conserved in all iPGAMs. Those labelled in green are from the phosphatase domain, and those in purple from the transferase domain. Ser75 becomes phosphorylated during the catalytic mechanism. Metal site 1 (M1) shows high occupancy even at low concentrations of Co^{2+} (10 μM), but metal site 2 (M2) is only partially occupied under these conditions. Site M1 adopts octahedral geometry, whereas site M2 is tetrahedral.

A main goal of our research is to identify new drugs for the treatment of neglected diseases such as human African trypanosomiasis, Chagas disease and a range of leishmaniases which are all caused by protists belonging to the family Trypanosomatidae. Together, these parasites cause severe public health and economic burdens on populations that are often already caught in a tragic cycle of poverty, poor nutrition and disease. As indicated previously, iPGAM is an attractive drug target, but for high-throughput and virtual screening, as well as for subsequent medicinal chemistry strategies it is essential to be able to understand and manipulate the conformation of the enzyme. This is particularly true for iPGAM with its small active site, so that conditions may be identified that favour the open conformation. Metals are likely to be essential players in this work.

We now report the metal analyses of cytosolic fractions from *T. brucei*, and of purified bacterially expressed

Lm iPGAM samples by ICP-OES and by ICP-MS, respectively. The results indicate that Zn^{2+} and Mg^{2+} instead of Co^{2+} are the biologically relevant metals, and that the cellular activity of iPGAM is only a small fraction of its potential activity.

Experimental

Preparation of *T. brucei* cytosolic fractions

T. brucei cytosolic fractions were prepared by differential centrifugation as described previously, in isotonic buffer (250 mM sucrose (UCB S.A., Brussels, Belgium), 25 mM Tris-HCl (Fisher Scientific), pH 7.4, 1 mM EDTA (Fluka Biochemika)).[28] EDTA was included primarily as a proteolytic inhibitor, but is also a useful form of metal sequestration. It was not removed from the fractions, and would thus still be present with its bound metals during the ICP-OES measurements. Soluble proteins in the cytosolic fractions were obtained in the supernatant after centrifugation at 50 000×*g*. Seven samples (3 mL each) were obtained that had been prepared on different occasions: three cytosolic samples from *in vitro* cultured procyclic insect-stage parasites (S1, S2, S3), and four cytosolic samples from pathogenic bloodstream-form parasites (S4, S5, S6 and S7). Samples S5 and S6 gave metal analysis results with relatively high standard deviations, and have been omitted.

Protein determination

Concentrations of proteins in cytosolic samples were measured by Bradford assay,[29] with Coomassie Plus Protein Assay Reagent supplied by Thermo Scientific, and with purified *Lm* iPGAM as a standard. Concentrations of purified *Lm* iPGAM were determined by UV absorbance at A_{280} nm using a Jasco V-550 UV-VIS spectrophotometer using the standard Beer–Lambert law, with known extinction coefficient (42 080 M^{-1} cm^{-1}) and molecular mass (60 723 Da).

Enzyme assays

iPGAM activity was measured by monitoring the conversion of 3-phosphoglycerate (3-PGA) to 2-phosphoglycerate (2-PGA) coupled with the oxidation of NADH to NAD^+ by lactate dehydrogenase (LDH) *via* enolase and pyruvate kinase. Assays of each cytosolic fraction were done using a Multimode Plate Reader-Molecular Devices M5 instrument, by the addition of a 10 μL sample of the cytosolic fraction to give a final reaction mixture of 100 μL containing 100 mM triethanolamine (TEA)-HCl buffer pH 7.6 (Sigma Life Sciences), 1.5 mM 3-PGA (Sigma-Aldrich), 5 mM $MgCl_2$ (Fisher Scientific), 50 mM KCl (Fisher Scientific), 0.8 mM NADH (Roche), 1 mM ADP (Sigma-Aldrich), 2 units of enolase (Sigma-Aldrich), 4 units of pyruvate kinase (Sigma-Aldrich) and 6 units of LDH (Fluka Analytical). The decreased absorbance of NADH at A_{340} nm was used to obtain the rate of reaction for specific activity measurements (one unit corresponds to the conversion of 1 μmol of substrate min^{-1} mg^{-1} protein under standard conditions). Assays done during the purification of *Lm* iPGAM used 10 μL samples containing ~0.2 mg mL^{-1} iPGAM which were added into a cuvette to give a final reaction mixture of 1000 μL containing 100 mM triethanolamine (TEA)-HCl buffer, pH 7.6, 1.5 mM 3-PGA, 5 mM $MgCl_2$, 50 mM KCl, 0.2 mM NADH, 1 mM ADP, 2 units of enolase, 4 units of pyruvate kinase and 6 units of LDH. The decrease in A_{340} nm was monitored using a Jasco V-550 UV-VIS spectrophotometer.

iPGAM expression and purification

Lm iPGAM with a C-terminal His_6 tag was expressed in *Escherichia coli* BL21(DE3) cells, and cell pellets were collected by centrifugation, with storage at -80 °C as described previously.[30] Cell lysis was done with a Constant Systems TS Cell Disruptor set to a maximum pressure of 22 000 psi in buffer A (50 mM NaH_2PO_4 (AnalaR BDH), 300 mM NaCl (Fisher Scientific) pH 8) with the addition of EDTA-free Protease Inhibitor Mixture (Roche) tablets (1 tablet per 25 mL of extract). The extract was centrifuged for 45 min at 4 °C at 43 400×*g*, and the supernatant injected onto an immobilised metal affinity chromatography column (FastFlow IMAC, GE Healthcare) pre-charged with 0.1 M Ni^{2+} (nickel(II) sulphate hexahydrate, Sigma-Aldrich), and washed with water, then buffer A in an ÄKTA HPLC system maintained at 10 °C. Following sample injection, the column was washed with buffer A, followed by 20 column volumes of buffer A containing 25 mM imidazole (Fisher Scientific). The *Lm* iPGAM was then eluted with a step of buffer A containing 500 mM imidazole, and desalted and buffer exchanged on a HiPrep 26/10 column (GE Healthcare) eluted with buffer B (20 mM HEPES (Sigma Life Sciences) pH 7.6). The protein was then buffer exchanged into 20 mM TEA-HCl, pH 7.6 and 50 mM NaCl using a PD10 column (GE Healthcare), and concentrated using a Vivaspin concentrator (10 000 MW cut off) (GE Healthcare) to ~2 mg mL^{-1}. The protein was stored in the cold room (4 °C) for further biochemical and biophysical assays or at −20 °C with the presence of 10% glycerol.

ICP-OES

ICP-OES was used to analyse cytosolic fractions with a Perkin Elmer Optima 5300 DV instrument, using a radio frequency (RF) forward power of 1400 W, with argon gas flows of 15, 0.2 and 0.75 L min^{-1} for plasma, auxiliary, and nebuliser flows, respectively. Using a peristaltic pump, the samples were taken up into a Gem Tip cross-Flow nebuliser and Scotts spray chamber at a rate of 1.50 mL min^{-1}. The instrument was operated in axial mode, and the selected wavelengths for each element were analysed in fully quantitative mode (three points per unit wavelength). Three replicate runs per sample were carried out. Initially two wavelengths were selected for each element, but after the analysis was completed the following wavelengths were chosen for reporting results: Co 228.616 nm, Cu 327.393 nm, Fe 238.204 nm, Mg 285.213 nm, Mn 257.610 nm, Ni 231.604 nm and Zn 206.200 nm.

To assess the limit of detection (LOD) for each metal and the matrix effects of the buffers, a range of calibration standards was prepared both in deionised water (18Ω, Elga USF) and in EDTA-free cytosol buffer (250 mM sucrose (Fisher Scientific), 25 mM Tris-HCl (Sigma Life Sciences), pH 7.4). Standards of 0, 0.1, 10 and 100 mg L^{-1} were prepared by sequential dilution from single element 1000 mg L^{-1} stock solutions (Fisher Scientific). With all of the calibration lines the correlation coefficients for the linear regression were

0.9992 or better. Ten blank samples of each buffer were analysed after the instrument was calibrated to calculate the limits of detection for the analysis. LOD = 2.26*SD_b/Slope (where SD_b is the standard deviation of the 10 blanks and 2.26 is the student t-test at the 95% confidence limit). The limits of detection are indicated in Fig. 3.

For analysis of the samples, a range of calibration standards from 0–500 mg L^{-1} was prepared in cytosol buffer. The majority of the samples fell in the range 200–500 mg L^{-1}, and results were calculated against a calibration from 10–500 mg L^{-1}. However where values were lower than 20 mg L^{-1}, the samples were reintegrated against a calibration only extending to 10 mg L^{-1}. Where values fell below the limit of detection they were quoted as a < value.

It is relevant to mention that attempts were initially made to analyse the cytosolic fractions by ICP-MS, but were found to be unsuitable because of technical problems.

ICP-MS

Purified protein samples were analysed using an Agilent 7500ce (with octopole reaction system) instrument, employing an RF forward power of 1540 W and reflected power of 1 W, with argon gas flows of 0.82 L min^{-1} and 0.2 L min^{-1} for carrier and makeup flows, respectively. For the collision cell the helium flow was 0.6 L min^{-1}. Sample solutions were taken up into the Micro mist using a peristaltic pump at a rate of approximately 1.2 mL min^{-1}. Skimmer and sample cones were made of nickel. The instrument was operated in spectrum multi-tune acquisition mode and five replicate runs per sample were employed. Each mass was analysed in fully quantitiative mode (three points per unit mass). The following masses were selected for analysis: ^{55}Mn, ^{59}Co and ^{63}Cu which were analysed in 'standard gas mode', and ^{26}Mg, ^{56}Fe, ^{60}Ni and ^{66}Zn which were analysed in 'helium gas mode' using the collision cell to remove any polyatomic interferences.

Calibration standards with a range of metal concentrations (0, 1, 10 and 100 μg L^{-1}) were prepared by sequential dilution from single element 1000 mg L^{-1} stock solutions (Fisher Scientific) in 20 mM TEA buffer pH 7.6 and 50 mM NaCl, as well as in water to assess matrix effects. Samples were analysed using the standards prepared in 20 mM TEA buffer pH 7.6. Ten blank samples of the 20 mM TEA buffer pH 7.6 buffer were analysed to calculate the LOD for each metal, and are indicated in Fig. 5. Where values fell below the limit of detection they were quoted as a < value.

Bioinformatics

Protein structure diagrams were generated using the program PyMOL.[31] The calculated extinction coefficient and molecular mass of *Lm* iPGAM were obtained from the Expasy ProtParam website (http://web.expasy.org/protparam/).

Results and discussion

Preparation and characterisation of *Tb* iPGAM cytosolic fractions

Trypanosomatids possess a unique subcellular organelle, the glycosome, with multiple copies in each cell which sequester the first seven of the ten glycolytic enzymes. iPGAM however, is located in the cytosol[8,20,32] along with enolase and pyruvate kinase. Cytosolic samples were prepared at the Université catholique de Louvain, Brussels, Belgium from both procyclic and bloodstream stages of *T. brucei*. The former stage occurs in the tsetse fly vector, whereas the latter stage corresponds to the pathogenic form found in mammalian blood. The cytosolic fractions were analysed by SDS-PAGE (Fig. 2), and gave sharp bands indicative of the absence of proteolytic degradation. The patterns of bands are broadly similar across all the fractions, with the exception of a prominent band with a mass of ~50 kDa in sample S4. The identity of this protein is not known.

The protein concentrations of the fractions showed variation (*i.e.* from 2.8 to 8.1 mg mL^{-1}), so the results for specific activity and for metal contents were all normalised to 1 mg of protein. The protein concentrations, iPGAM specific activities and estimated iPGAM concentrations of the fractions are given in Table 1. The specific activities were measured for the cytosolic fractions with no additional Co^{2+}. It can be seen that the specific activities of iPGAM in the procyclic cytosolic fractions were higher than in the bloodstream-form cytosolic fractions. As anticipated, the addition of Co^{2+} gave a substantial enhancement of activity (about 10-fold, data not shown).

ICP-OES analysis of *T. brucei* cytosol shows the presence of Mg^{2+}, Zn^{2+} and Fe^{2+}, and absence of Co^{2+}

The concentrations of seven different metals (Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}) were measured by ICP-OES analysis of cytosolic fractions from procyclic- and bloodstream-form *T. brucei*. This type of analysis gives the total metal concentrations, and does not distinguish between metals bound to macromolecules or to EDTA, and those unbound. Elemental standards were prepared as described in the Experimental section, and gave linear concentration-response plots in both water and in cytosol buffer. Any background readings or noise signals collected from the control data were deducted from the readings of the samples. Mg^{2+}, Zn^{2+} and Fe^{2+} were found to be the most abundant metals (~200, 30 and 15 μM, respectively; Fig. 3), with Cu^{2+}, Mn^{2+} and Ni^{2+} below 1 μM,

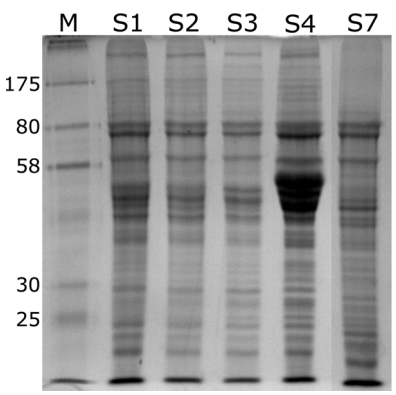


Fig. 2 Electrophoretic analysis of cytosolic fractions. A 12% SDS-PAGE gel was used with 5 μL samples of each cytosolic fraction. The masses (in kDa) of standard marker proteins are indicated in lane M. Abbreviations: M, marker proteins; S1–S3, cytosolic fractions from procyclic *T. brucei*; S4 and S7, cytosolic fractions from bloodstream-form *T. brucei*.

Table 1 iPGAM molar concentration in *T. brucei* cytosolic fractions

	S1[a]	S2	S3	S4	S7
Protein concentration (mg mL^{-1})	4.0	3.9	2.8	8.1	5.4
Specific activity (nmol min^{-1} mg^{-1} protein)	16.0	19.7	6.29	1.12	5.80
iPGAM concentration in cytosolic fractions (μmoles L^{-1})[b]	50.1	60.3	13.8	7.12	24.5

[a] Samples: S1, S2 and S3 are procyclic fractions; S4 and S7 are from bloodstream parasites. [b] The calculation was made using the assumption that pure iPGAM has a specific activity of 21 μmole min^{-1} mg^{-1} protein (under the same assay conditions as those used to measure the activity of iPGAM in the cytosol).

and Co^{2+} below the limit of detection (<0.035 μM). It can be seen that the samples from the two different life stages of *T. brucei* gave similar results.

The specific activity of bacterially expressed *Tb* iPGAM purified in the absence of Co^{2+} is ~25 μmol min^{-1} mg^{-1} protein,[8,19] very similar to the value of 21 μmol min^{-1} mg^{-1} protein determined for pure bacterially expressed *Lm* iPGAM in this study (see next section). The cytosolic concentrations of *Tb* iPGAM may be estimated from the specific activity of the pure enzyme compared to the specific activities within the cytosolic fractions, and give a value of approximately 10-50 μM (Table 1). It is clear from the ICP-OES results that the Co^{2+} concentrations in the cytosolic fractions are <0.035 μM, and are therefore insufficient to satisfy even a small proportion of the iPGAM present.

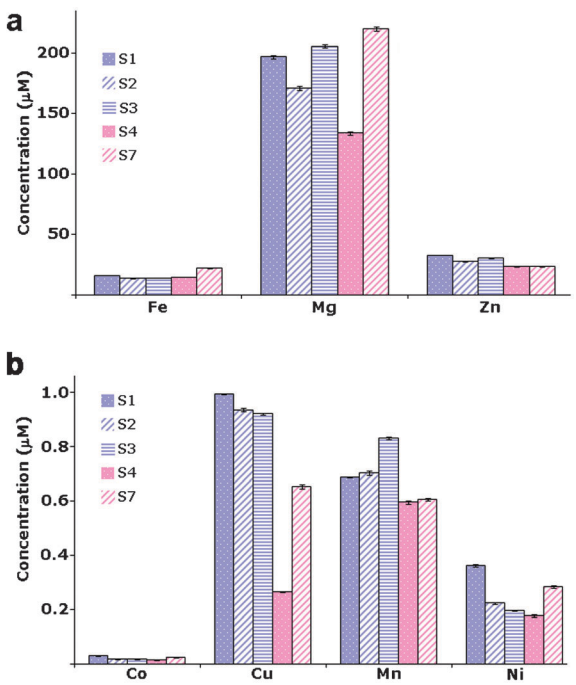


Fig. 3 ICP-OES of cytosolic fractions from procyclic and bloodstream-form *T. brucei*. The concentrations of metals in the cytosolic fractions are shown. The blue bars correspond to the procyclic cytosolic fractions S1, S2 and S3, whereas the pink bars indicate bloodstream cytosolic fractions S4 and S7. (a) Results for the relatively abundant metals, Fe^{2+}, Mg^{2+} and Zn^{2+}, with the *y* axis scale 0–200 μM. (b) Results for the less abundant metals, Co^{2+}, Cu^{2+}, Mn^{2+} and Ni^{2+}, with the *y* axis scale 0–1.0 μM. The values for Co (0.0130–0.0296 μM) are below the limit of detection (<0.0352 μM). The following limit of detection values (all μM) were obtained (see Experimental section): Fe, 0.0819; Mg, 0.495; Zn, 0.122; Co, 0.0352; Cu, 0.116; Mn, 0.0920; Ni, 0.104.

Purification and characterisation of *Lm* iPGAM

The bacterial expression system for *Lm* iPGAM was the same as described previously.[20] In the present study however, purified *Lm* iPGAM was obtained in the absence of added Co^{2+} from a two-step purification procedure: IMAC chromatography and desalting with buffer exchange (Fig. 4). The Ni^{2+}-containing IMAC column in this procedure replaced the previously used Co^{2+}-containing Talon (Clontech) column,[30] and the 10 μM $CoCl_2$ used throughout previous purification procedures was omitted. The yield of protein obtained from 1 L of cell culture was found to be ~20 mg. SDS-PAGE analysis showed that the purified *Lm* iPGAM gave a single prominent band with mass ~60 700 Da, corresponding to the expected mass (Fig. 4). The specific activity was calculated to be 21 μmol min^{-1} mg^{-1} protein for the conversion of 3-PGA to 2-PGA, similar to the value of 26 μmol min^{-1} mg^{-1} protein obtained for pure bacterially expressed *Tb* iPGAM that had also been purified on a Ni^{2+} IMAC column without the addition of any divalent metals.[19] *Lm* iPGAM purified in the Edinburgh laboratory from a GE Healthcare IMAC column pre-charged with 0.1 M Co^{2+} or from a Talon column had specific activities of 370 ± 2 μmol min^{-1} mg^{-1} or ~400 μmol min^{-1} mg^{-1} protein,[33] respectively. These values correspond closely with *Lm* iPGAM purified from a Talon column in the Brussels laboratory with a specific activity of 419 ± 4 μmol min^{-1} mg^{-1}.[20] The activity assays for all these enzymes were carried out in the presence of Co^{2+}, and clearly all gave an approximately 20-fold enhancement of activity over *Lm* iPGAM purified and assayed in the absence of Co^{2+}.

ICP-MS analysis of purified *Lm* iPGAM fails to detect Co^{2+}

The metal content of the essentially homogenous purified bacterially expressed *Lm* iPGAM was analysed by the more sensitive ICP-MS technique. As for ICP-OES, the concentrations of seven different metals (Co^{2+}, Cu^{2+}, Fe^{2+}, Mg^{2+}, Mn^{2+}, Ni^{2+} and Zn^{2+}) were measured. Elemental standards were prepared as described in the Experimental section, and gave linear concentration-response curves in both water and in enzyme storage buffer (20 mM TEA-HCl, pH 7.6 and 50 mM NaCl). In the case of the purified enzyme, Mg^{2+} was the most abundant metal (as it was in the cytosolic fractions), but the amount of Zn^{2+} was substantially enhanced, and approached the concentration of Mg^{2+} (Fig. 5). The concentration of Co^{2+} was 0.01 μM, above the limit of detection (<0.0002 μM), but very low compared to the concentration of iPGAM (0.99 μM). Taken together, these results show clearly that despite Co^{2+} being the

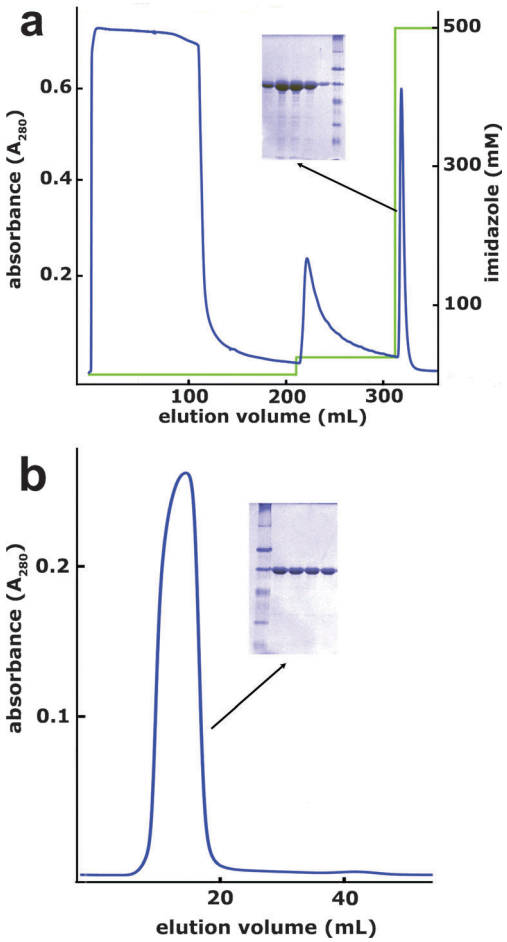


Fig. 4 Purification of *Lm* iPGAM. The two-step procedure involved (a) IMAC *via* the enzyme's C-terminal His tag on a Ni^{2+}-charged column, followed by (b) desalting and buffer exchange with a HiPrep 26/10 column. SDS-PAGE gels of peak fractions from both steps are shown. The standard protein markers are as in Fig. 2. The purified enzyme was homogenous with the expected mass of 60 700 Da, as shown by the SDS-PAGE gel in panel b. Blue traces correspond to absorbance at 280 nm, and the green trace in panel a indicates the concentration of imidazole.

metal that enables *Tb* iPGAM and *Lm* iPGAM to be maximally active, it is not present in the enzyme in cytosolic fractions or in enzyme purified in the absence of added Co^{2+}. The molar ratios of Mg^{2+} and Zn^{2+} to *Lm* iPGAM in the purified enzyme were ~8 and ~6 μmoles of metal to 1 μmole of *Lm* iPGAM, respectively, indicating that either or both of these metals may be the preferred metal(s) *in vivo*. By contrast, Mn^{2+} that is known to be the preferred metal for bacterial iPGAMs[15,18] was present in the purified *Lm* iPGAM at a molar ratio of less than 0.3. It is noteworthy that the concentration of Ni^{2+} was low, despite the purification being done on an IMAC column charged with Ni^{2+}.

Mg^{2+} and Zn^{2+} are likely to be the biologically relevant metals of trypanosomatid iPGAMs

It is generally accepted that enzymes are evolving by natural selection to attain a state of optimal catalytic activity, often with adaptation to a particular biological environment. As an

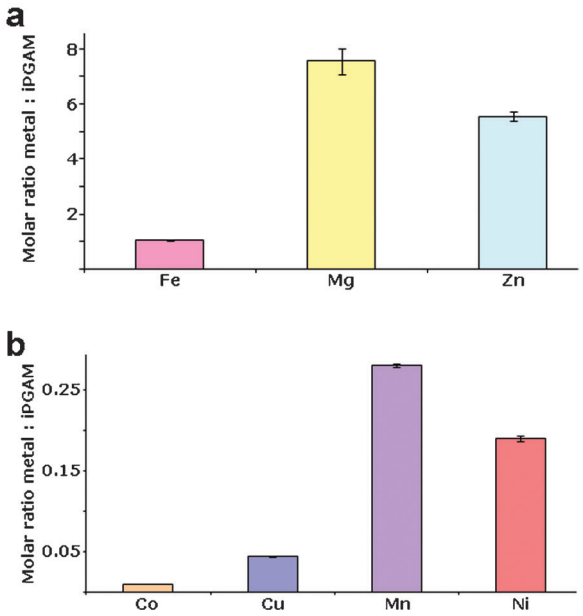


Fig. 5 ICP-MS analysis of purified *Lm* iPGAM. The relative molar concentrations of metals and purified *Lm* iPGAM are shown. (a) Abundant metals, Fe, Mg and Zn, with the *y* axis scale 0–8. (b) Low abundance metals, with the *y* axis scale 0–0.25. The following limit of detection values (all μM) were obtained (see Experimental section): Fe, 0.0313; Mg, 0.195; Zn, 0.00652; Co, 0.000238; Cu, 0.000936; Mn, 0.00128; Ni, 0.000803.

example, the activity of the glycolytic enzyme triose phosphate isomerase is limited only by diffusion of substrate onto, and product off, the enzyme, and is therefore thought to have evolved to almost catalytic perfection.[34] There is thus an understandable tendency to correlate the highest specific activity of an enzyme *in vitro* with its 'native' state *in vivo*. This is clearly not the case for trypanosomatid iPGAMs where the optimal activity (by a factor of ten or more) is obtained with Co^{2+}, but the concentration of this metal is essentially undetectable in the parasite cytosol. In this respect, iPGAM is not alone, and carboxypeptidase A,[35] aminopeptidase,[4] alcohol dehydrogenase,[36] 5′-nucleotidase[37] and catechol dioxygenase[38] are just some among many examples of Zn^{2+}-containing enzymes *in vivo* that are hyperactive when Co^{2+} replaces the Zn^{2+} *in vitro*.

The results of ICP-OES indicate that only Mg^{2+} and Zn^{2+} are sufficiently abundant in the parasite cytosol to support the activity of a significant proportion of the iPGAM present (Fig. 3 and Table 1), whereas Cu^{2+}, Mn^{2+} and Ni^{2+} were all well below the concentration of iPGAM. A similar trend is observed with the ICP-MS analysis of purified *Lm* iPGAM in which only Mg^{2+} and Zn^{2+} are sufficiently abundant to support the activity of the enzyme (Fig. 5). It thus seems that the biologically relevant metal for this enzyme is likely to be either Mg^{2+} or Zn^{2+}, or both. As mentioned previously, iPGAM belongs to the metal-dependent alkaline phosphatase superfamily[10] most of whose members require two divalent metals at the active site. Four enzymes are particularly well characterised structurally: *E. coli* alkaline phosphatase,[22] *Bs* iPGAM (closed conformation),[15] *Ba* iPGAM (open conformation)[17] and *Lm* iPGAM.[16] Alkaline phosphatase uses

two Zn^{2+} ions, *Bs* iPGAM two Mn^{2+} ions, but *Lm* iPGAM by contrast, appears to function with a one-metal mechanism. Of the two metal sites of *Lm* iPGAM (Fig. 1), only site M1 shows high occupancy, with site M2 having only low occupancy (6%) even at 4 mM Co^{2+}. Moreover, when site M2 is occupied, the side chain of Ser75 that is phosphorylated during catalysis adopts a position unfavourable for phospho transfer.[16] Site M1 has octahedral geometry, indicating that it has a preference for Mg^{2+} and Co^{2+} (MESPEUS database, http://mespeus.bch.ed.ac.uk/MESPEUS/_1.jsp). The poorly occupied site M2 has tetrahedral geometry with a likely preference for Zn^{2+}. It is not possible at this stage to decide definitively which metal is the biologically relevant one, and indeed it may be that both metals can support trypanosomatid iPGAM activity.

On a note of caution, it can be mentioned that the discovery that *Tb* iPGAM and *Lm* iPGAM are optimally active with Co^{2+} was an accident, and a direct consequence of two experimental developments: (i) the use of His tags as a very convenient purification method for bacterially expressed proteins, and (ii) the availability of the Co-containing Talon IMAC resin with enhanced affinity for His tags. Both innovations were adopted by the Brussels and Edinburgh research groups, and have been very efficient for the purification of large quantities of several trypanosomatid glycolytic enzymes, including iPGAMs. Furthermore, routine enzyme stability studies were carried out on *Tb* iPGAM and *Lm* iPGAM as a strategy to improve the outcome of crystallisation trials.[33] These studies showed that Co^{2+} was an excellent stabilisation agent for the trypanosomatid iPGAMs, and 10 μM $CoCl_2$ was thereafter frequently included in buffers.

An important corollary of the findings in this paper is that the specific activity of trypanosomatid iPGAM *in vivo* is relatively low compared with the specific activities of other glycolytic enzymes.[9] This is thus another reason why it is an attractive drug target because it exerts a higher flux control than most of the other glycolytic enzymes under typical metabolic conditions. The only exception is glyceraldehyde 3-phosphate dehydrogenase, which has a similar flux control coefficient[9] when measured under the same metabolic conditions. Moreover, as previously mentioned, there is essentially no problem about species specificity of inhibitors because human and parasite PGAMs share no similarity in structure or mechanism.

Conclusion

ICP-OES was used to determine the divalent metal contents of cytosolic fractions from procyclic and bloodstream-form *T. brucei*, and now provides for the first time essential information relevant to understanding the *in vivo* properties of trypanosomatid metal-dependent enzymes. Of particular interest is the metal dependency of iPGAM, which has previously been shown to be maximally active in the presence of Co^{2+}. The ICP-OES results show clearly that the biologically relevant metal for this enzyme cannot be Co^{2+} which was undetectable in the cytosolic fractions. It is likely that iPGAM is in fact a relatively inefficient Mg^{2+} and/or Zn^{2+}-dependent enzyme. The significance of these observations is being further explored with the development of a new discontinuous plate-reader activity assay coupled to enolase, pyruvate kinase and lactate dehydrogenase to be carried out with minimal metal interference of the latter enzymes. An understanding of the *in vivo* metal requirements of iPGAM will be important for the design and implementation of a drug discovery programme currently in progress.

Acknowledgements

We thank Melisa Gualdron-Lopez and Muriel Mazet, de Duve Institute, Brussels, Belgium for providing the cytosolic fractions, and Dr Buabarn Kuaprasert (Siam Synchrotron Light Research Institute) for helpful discussions. This project has been funded by the Ministry of Higher Education Malaysia and the University of Science Malaysia. The Centre for Translational and Chemical Biology and the Edinburgh Protein Production Facility were funded by the Wellcome Trust and the BBSRC. We also acknowledge the use of the Joint Chemistry/Geosciences ICP Facility, University of Edinburgh.

References

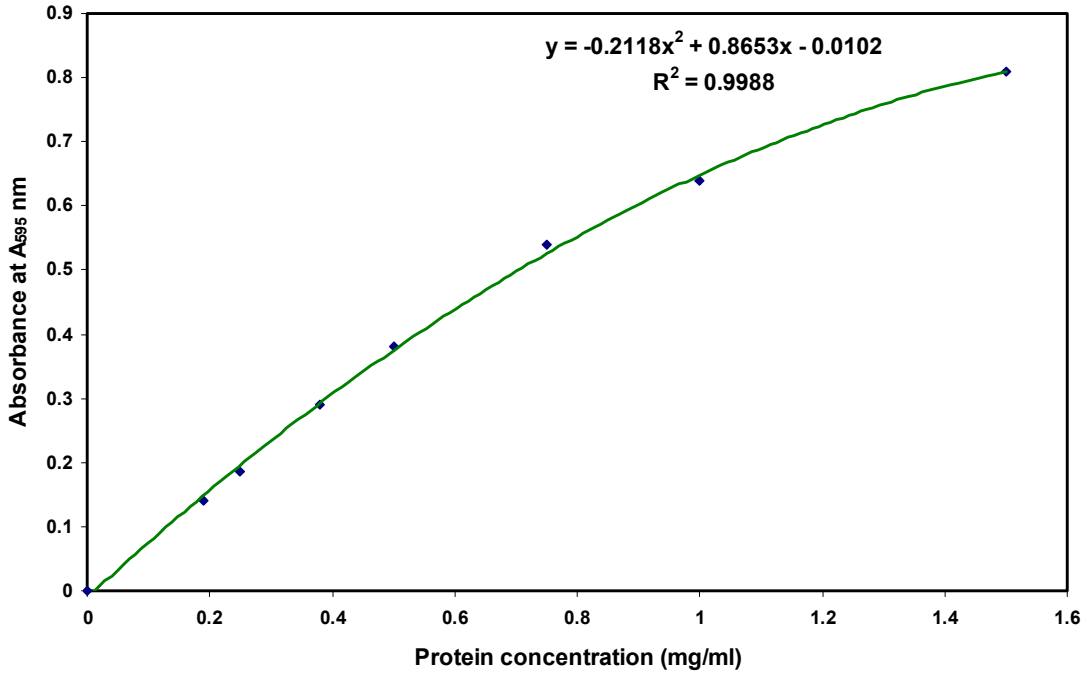
1 M. M. Harding, M. W. Nowicki and M. D. Walkinshaw, *Crystallogr. Rev.*, 2010, **16**, 1–56.
2 W. Maret, *Metallomics*, 2010, **2**, 117–125.
3 A. Cvetkovic, A. L. Menon, M. P. Thorgersen, J. W. Scott, F. L. Poole, F. E. Jenney, W. A. Lancaster, J. L. Praissman, S. Shanmukh, B. J. Vaccaro, S. A. Trauger, E. Kalisiak, J. V. Apon, G. Siuzdak, S. M. Yannone, L. A. Tainer and M. W. W. Adams, *Nature*, 2010, 779–782.
4 J. M. Prescott, F. W. Wagner, B. Holmquist and B. L. Vallee, *Biochemistry*, 1985, **24**, 5350–5356.
5 J. R. Dean, S. Munro, L. Ebdon, H. M. Crews and R. C. Massey, *J. Anal. At. Spectrom.*, 1987, **2**, 607–610.
6 D. J. Gray, R. Burns, P. Brunswick, J. Le Huray, W. Chan, M. Mast, K. Allamneni and S. Sreedharan, in *Plasma Source Mass Spectrometry: Current Trends and Future Developments*, ed. G. Holland and D. R. Bandura, Royal Society of Chemistry, 2005, pp. 43–58.
7 C. L. M. J. Verlinde, V. Hannaert, C. Blonski, M. Willson, J. Périé, L. A. Fothergill-Gilmore, F. R. Opperdoes, M. H. Gelb, W. G. J. Hol and P. A. M. Michels, *Drug Resist. Updates*, 2001, **4**, 1–14.
8 N. Chevalier, D. J. Rigden, J. Van Roy, F. Opperdoes and P. A. M. Michels, *Eur. J. Biochem.*, 2000, **267**, 1464–1472.
9 M.-A. Albert, J. R. Haanstra, V. Hannaert, J. Van Roy, F. R. Opperdoes, B. M. Bakker and P. A. M. Michels, *J. Biol. Chem.*, 2005, **280**, 28306–28315.
10 M. Y. Galperin, A. Bairoch and E. V. Koonin, *Protein Sci.*, 1998, **7**, 1829–1835.
11 L. A. Fothergill-Gilmore and P. A. M. Michels, *Prog. Biophys. Mol. Biol.*, 1993, **59**, 105–236.
12 G. C. Smith and L. F. Hass, *Biochem. Biophys. Res. Commun.*, 1985, **131**, 743–749.
13 G. C. Smith, A. D. McWilliams and L. F. Hass, *Biochem. Biophys. Res. Commun.*, 1986, **136**, 336–340.
14 C. M. Johnson, M. J. Gore and N. C. Price, *Biochem. Soc. Trans.*, 1987, **15**, 878.
15 M. J. Jedrzejas, M. Chander, P. Setlow and G. Krishnasamy, *EMBO J.*, 2000, **19**, 1419–1431.
16 M. W. Nowicki, B. Kuaprasert, I. W. McNae, H. P. Morgan, M. M. Harding, P. A. M. Michels, L. A. Fothergill-Gilmore and M. D. Walkinshaw, *J. Mol. Biol.*, 2009, **394**, 535–543.
17 M. Nukui, L. V. Mello, J. E. Littlejohn, B. Setlow, P. Setlow, K. Kim, T. Leighton and M. J. Jedrzejas, *Biophys. J.*, 2007, **92**, 977–988.
18 N. J. Kuhn, B. Setlow, P. Setlow, R. Cammack and R. Williams, *Arch. Biochem. Biophys.*, 1995, **320**, 35–42.
19 J.-F. Collet, V. Stroobant and E. Van Schaftinger, *FEMS Microbiol. Lett.*, 2001, **204**, 535–543.

20 D. G. Guerra, D. Vertommen, L. A. Fothergill-Gilmore, F. R. Opperdoes and P. A. M. Michels, *Eur. J. Biochem.*, 2004, **271**, 1798–1810.
21 S. Raverdy, Y. Zhang, J. Foster and C. K. Carlow, *Mol. Biochem. Parasitol.*, 2007, **156**, 210–216.
22 E. E. Kim and H. W. Wyckoff, *J. Mol. Biol.*, 1991, **218**, 449–464.
23 J. J. R. Fraústo da Silva and R. J. P. Williams, in *The Biological Chemistry of the Elements*, Oxford University Press, 2nd edn, 2001.
24 K. Håkansson and A. Wehnert, *J. Mol. Biol.*, 1992, **228**, 1212–1218.
25 I. Bertini, C. Luchinat, R. Pierattelli and A. J. Vila, *Eur. J. Biochem.*, 1992, **208**, 607–615.
26 S. Lindskog, *Pharmacol. Ther.*, 1997, **74**, 1–20.
27 B. Bennett, in *Metals in Biology: Applications of High-Resolution EPR to Metalloenzymes*, ed. G. Hanson and L. Berliner, Springer Science, 2010, ch. 10.
28 O. Misset and F. R. Opperdoes, *Eur. J. Biochem.*, 1984, **144**, 475–483.
29 M. M. Bradford, *Anal. Biochem.*, 1976, **72**, 248–254.
30 B. Poonperm, D. G. Guerra, I. W. McNae, L. A. Fothergill-Gilmore and M. D. Walkinshaw, *Acta Crystallogr., Sect. D: Biol. Crystallogr.*, 2003, **59**, 1313–1316.
31 W. L. DeLano, DeLano Scientific, San Carlos, CA, 2002.
32 K. K. Oduro, I. W. Flynn and I. B. R. Bowman, *Exp. Parasitol.*, 1980, **50**, 123–135.
33 B. Poonperm, *PhD Thesis*, University of Edinburgh, 2005.
34 W. J. Albery and J. R. Knowles, *Biochemistry*, 1976, **15**, 5631–5640.
35 W. D. Behnke and B. L. Vallee, *Proc. Natl. Acad. Sci. U. S. A.*, 1972, **69**, 2442–2445.
36 M. Cavaletto, E. Pessione, A. Vanni and C. Giunta, *J. Biotechnol.*, 2000, **84**, 87–91.
37 L. McMillen, I. R. Beacham and D. M. Burns, *Biochem. J.*, 2003, **372**, 625–630.
38 A. J. Fielding, E. G. Kovaleva, E. R. Farquhar, J. D. Lipscomb and L. Que, *J. Biol. Inorg. Chem.*, 2010.

Appendix I

Appendix I (a)

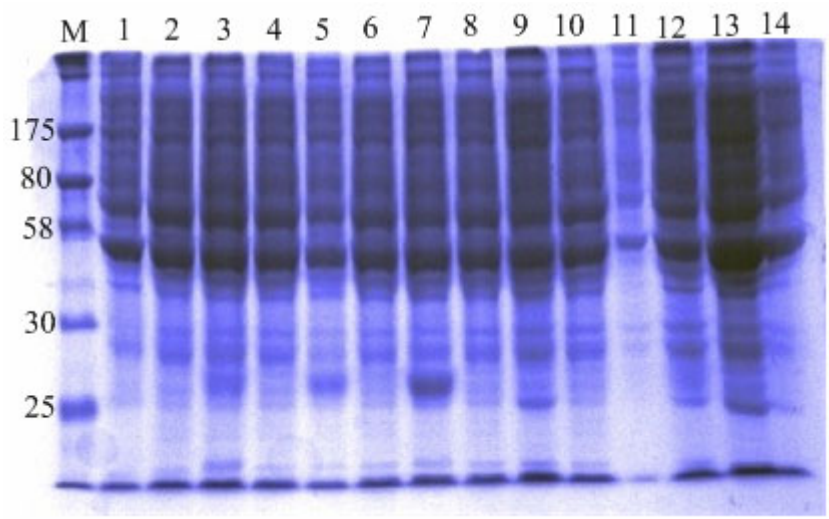
Bradford standard curve



The Bradford curve was constructed by using BSA in the range of 1.5, 1, 0.75, 0.5, 0.38, 0.25 and 0.19 mg/ml. BSA dilutions were carried out with a sucrose buffer (250 mM sucrose, 25 mM Tris-HCl pH 7.4). 20 µl of each concentration of BSA were mixed with 980 µl of the assay reagent (Coomassie Plus™ Protein Assay Reagent) (Thermo Scientific). The absorbance reading was at 595 nm.

Appendix I (b)

The un-tagged *Lm*iPGAM expression trial



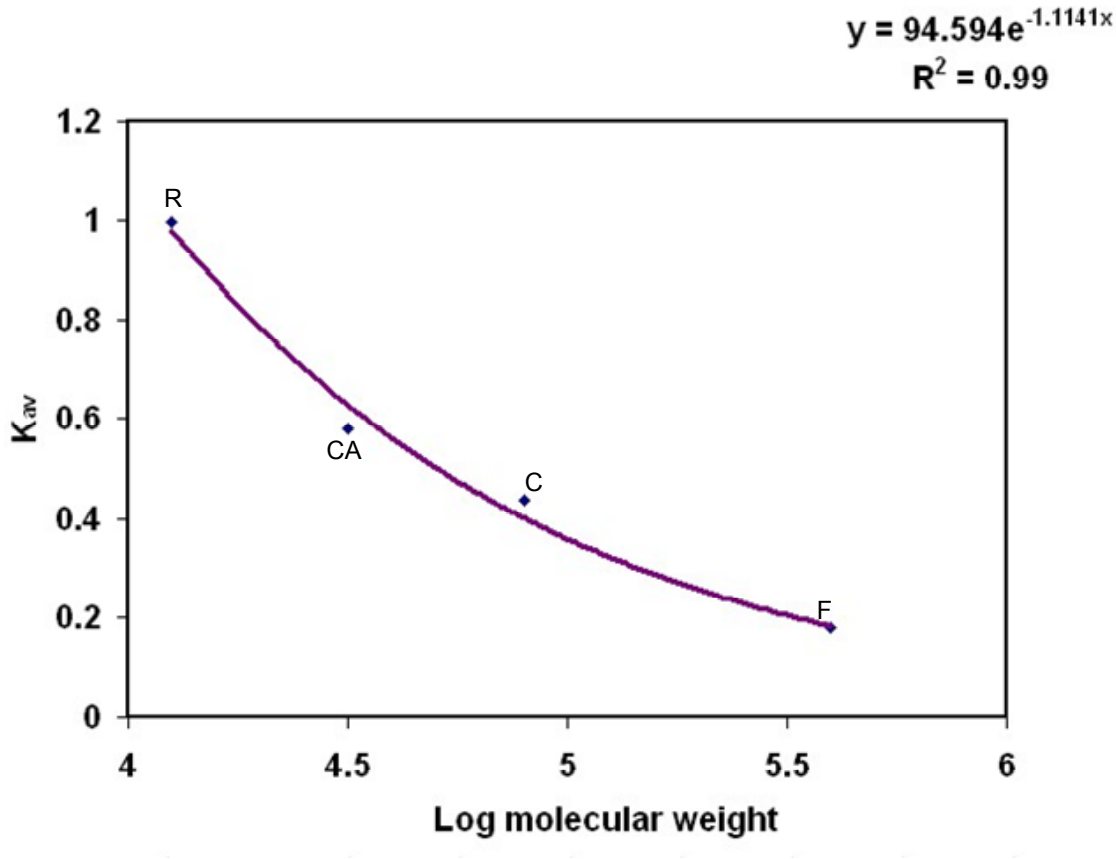
The un-tagged *Lm*iPGAM expression trials at 37°C. Wells 1,3,5,7,9,11,13 and 15 are the induced samples, whereas wells 2,4,6,8,10,12 and 14 are the un-induced samples (negative controls). The expression conditions for each sample are listed below:

Sample	Growth medium	Cell lines	IPTG concentration (mM)
1	LB	BL21 (DE3)	0.5
3	LB	BL21 (DE3) pLysS	0.5
5	LB	Rosetta (DE3) pLysS	0.5
7	LB	Rosetta (DE3) pLysS	1
9	Auto-Induction	BL21 (DE3)	Not required
11	Auto-Induction	BL21 (DE3) pLysS	Not required
13	Auto-Induction	Rosetta (DE3) pLysS	Not required
15	TB	BL21 (DE3)	1

Expression trials were also conducted at 20°C and 30°C. The results were observed to be very similar to this experiment and hence were not shown here.

Appendix I (c)

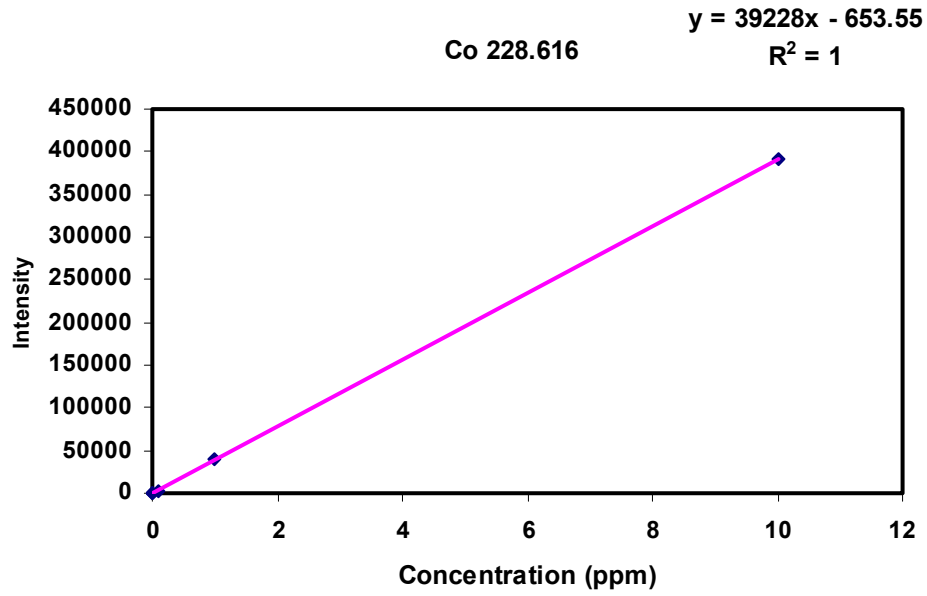
Gel filtration standard curve

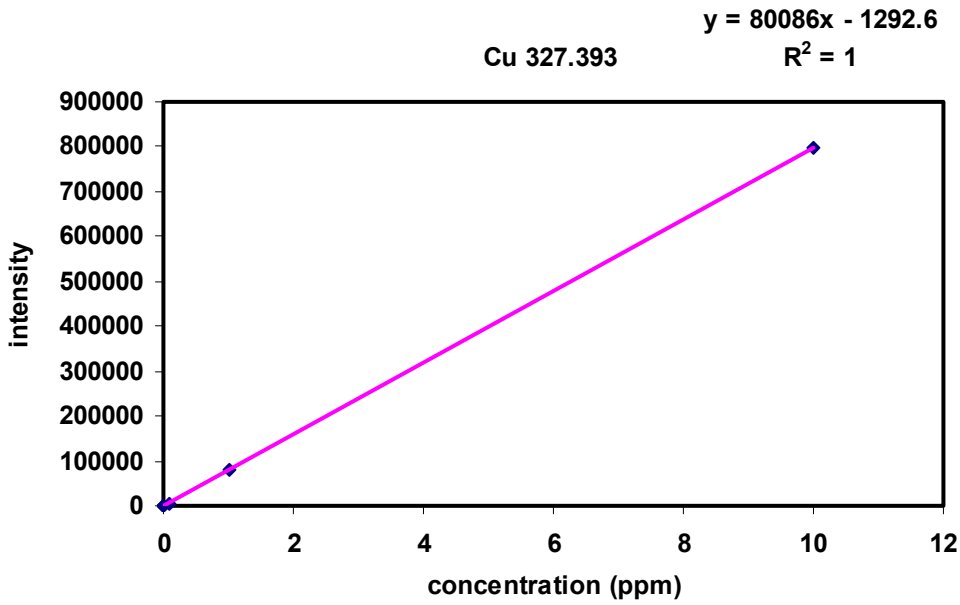


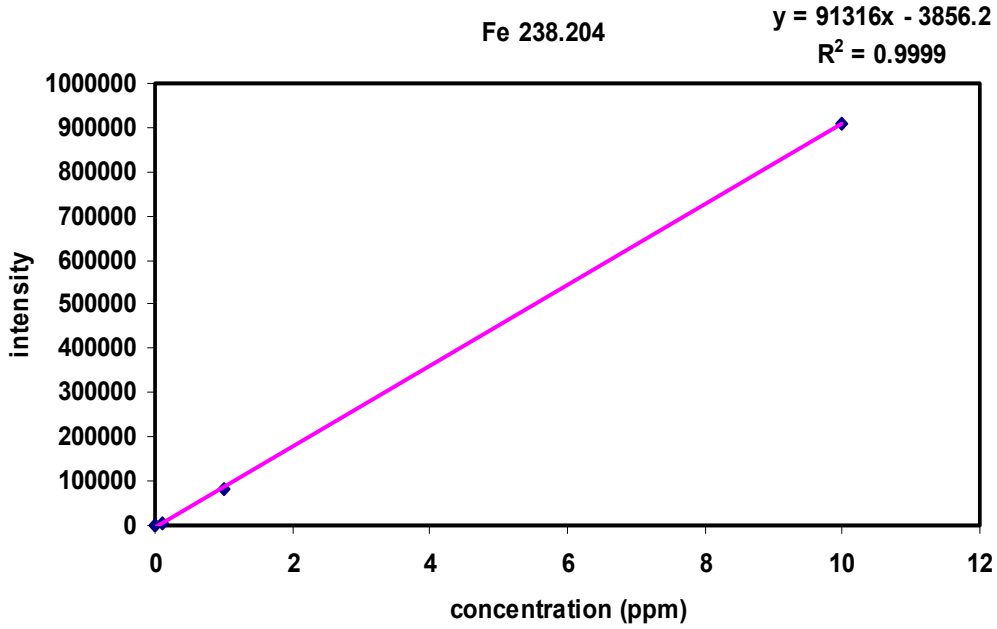
A standard curve of K_{av} versus log molecular weight (MW) for standard proteins (F = ferritin, C = conalbumin, CA = carbonic anhydrase and R = ribonuclease A) which were eluted from Superdex 200 10/300 column. The buffer contains 20 mM TEA pH 7.6 and 50 mM NaCl.

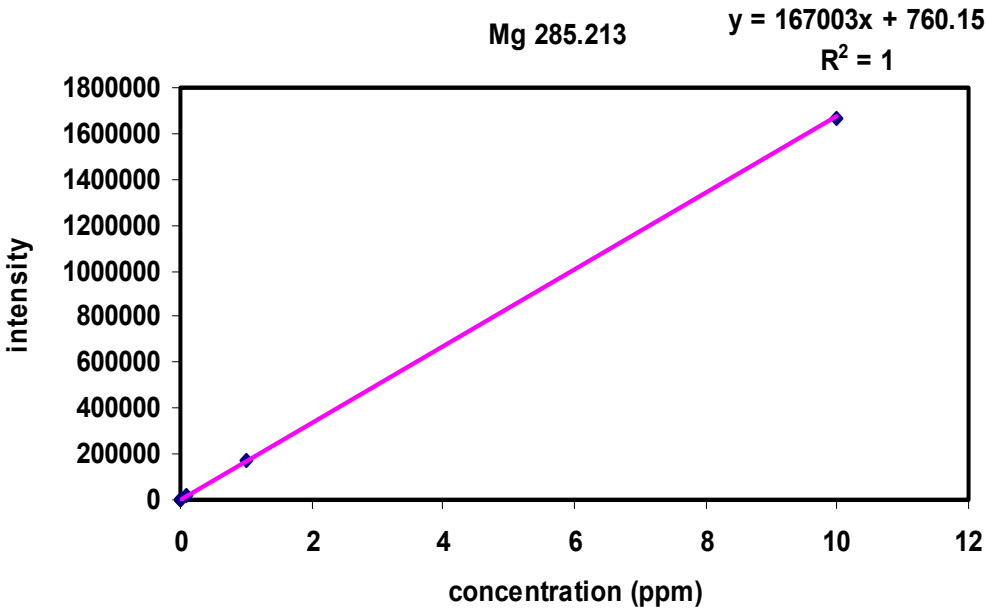
Appendix I (d)

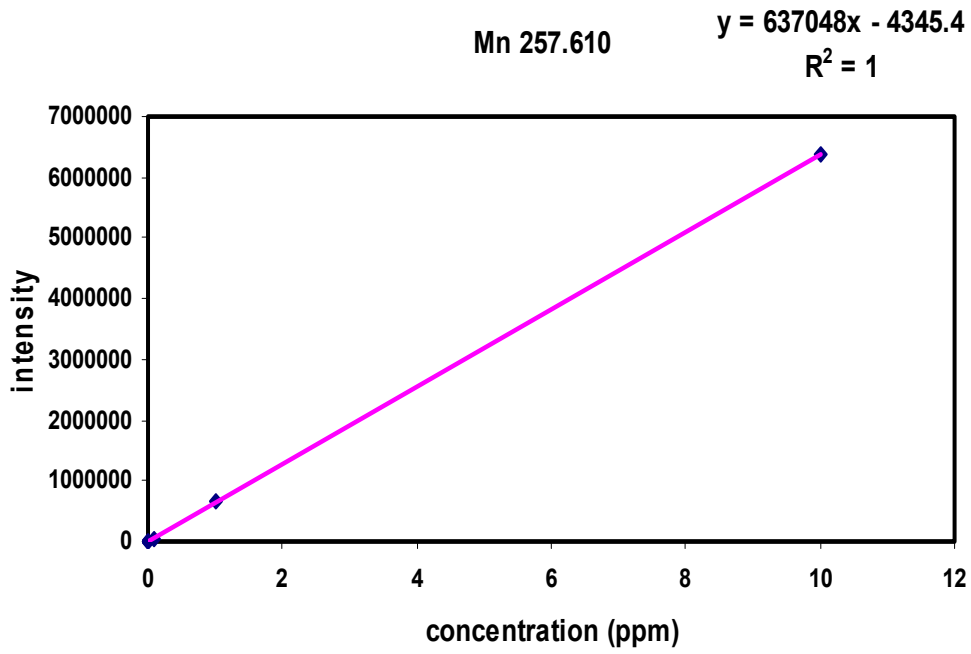
ICP-OES standard curves

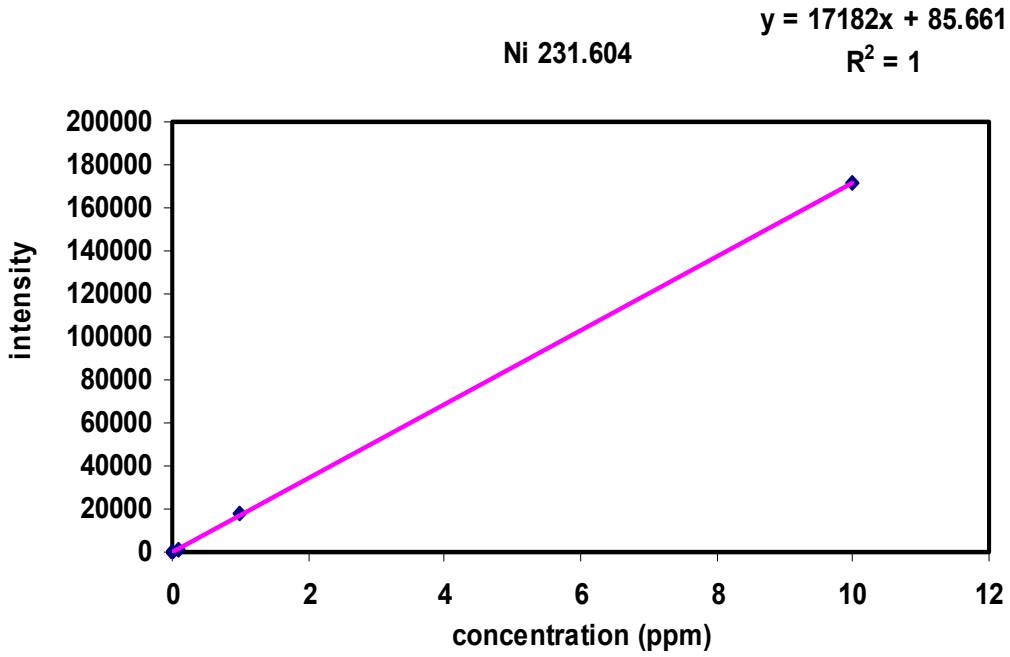


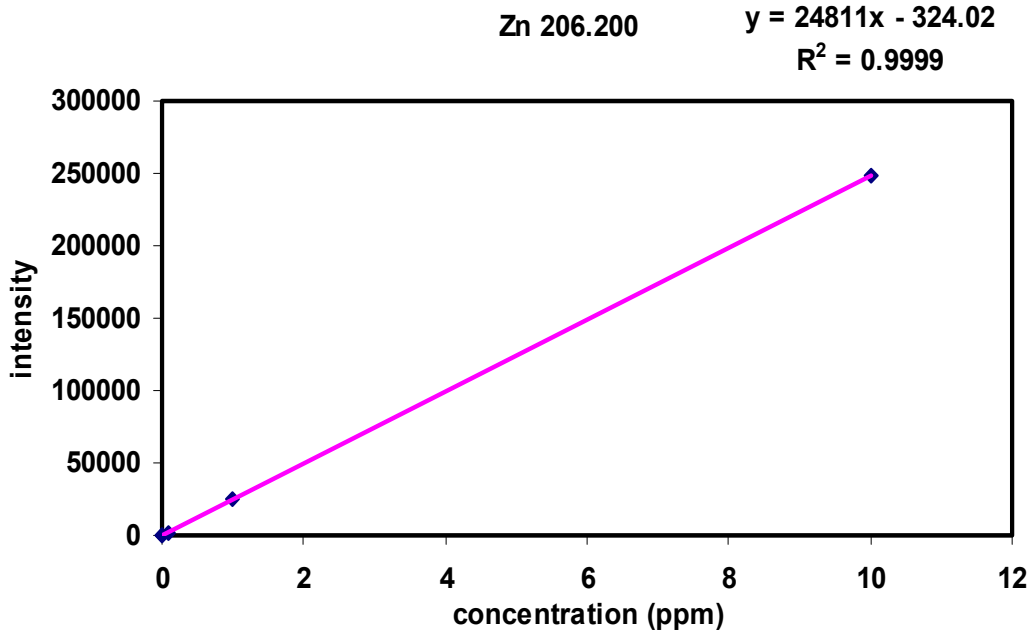



Fe 238.204
y = 91316x - 3856.2
R2 = 0.9999
1000000
900000
800000
700000
600000
500000
400000
300000
200000
100000
0
intensity
0
2
4
6
8
10
12
concentration (ppm)


Mg 285.213
y = 167003x + 760.15
R2 = 1
1800000
1600000
1400000
1200000
1000000
800000
600000
400000
200000
0
intensity
0
2
4
6
8
10
12
concentration (ppm)


Mn 257.610
y = 637048x - 4345.4
R2 = 1
intensity
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
0
2
4
6
8
10
12
concentration (ppm)

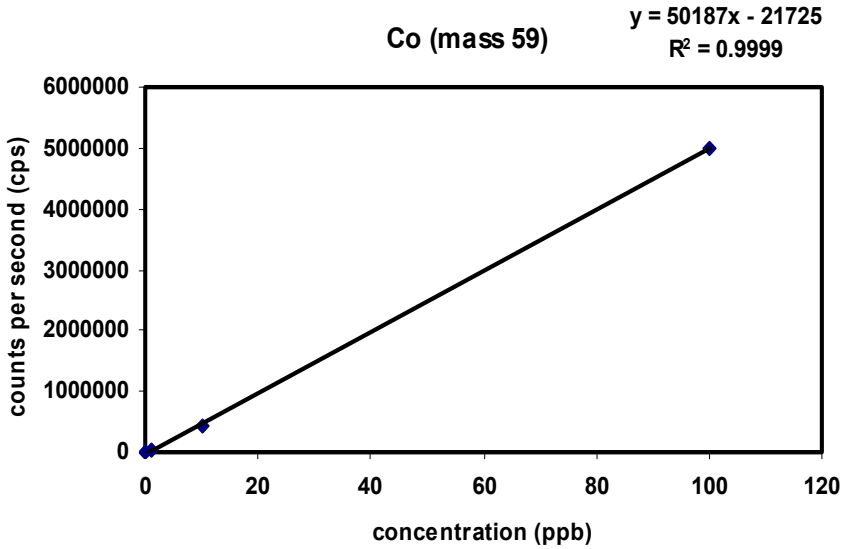

Ni 231.604
y = 17182x + 85.661
R2 = 1
intensity
200000
180000
160000
140000
120000
100000
80000
60000
40000
20000
0
0
2
4
6
8
10
12
concentration (ppm)

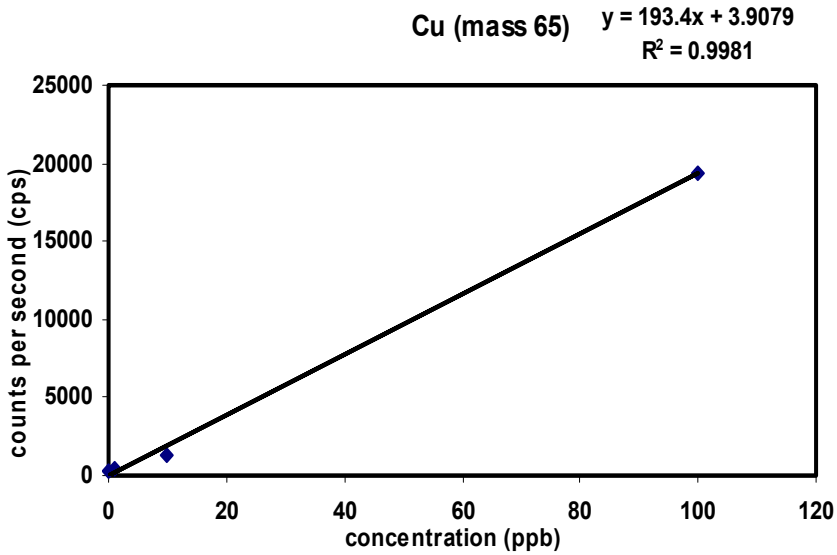


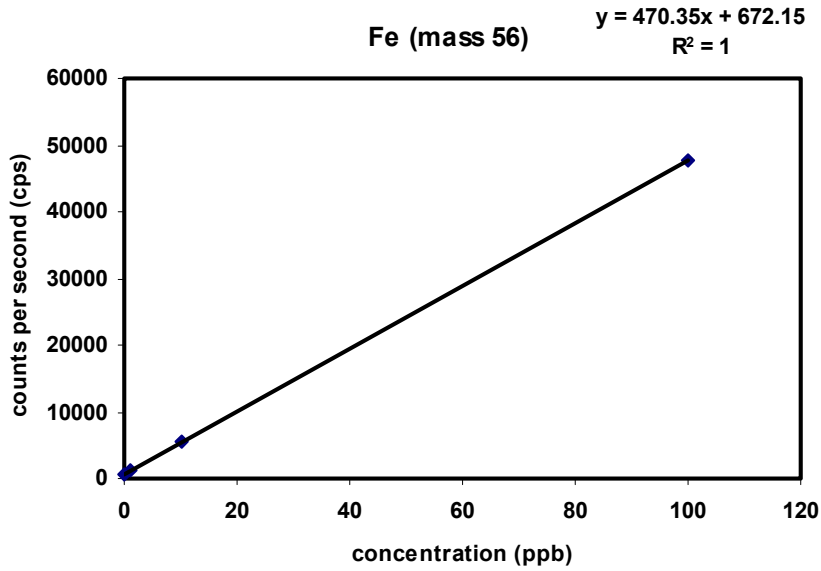
The calibration curves of concentration (in ppm or mg/L) versus intensity for ICP-OES analysis. The concentrations of divalent metals are in the range of 0.001, 0.01, 0.1, 1 and 10 ppm. The experiment was performed by using sucrose buffer (250 mM sucrose, 25 mM Tris-HCl pH 7.4).

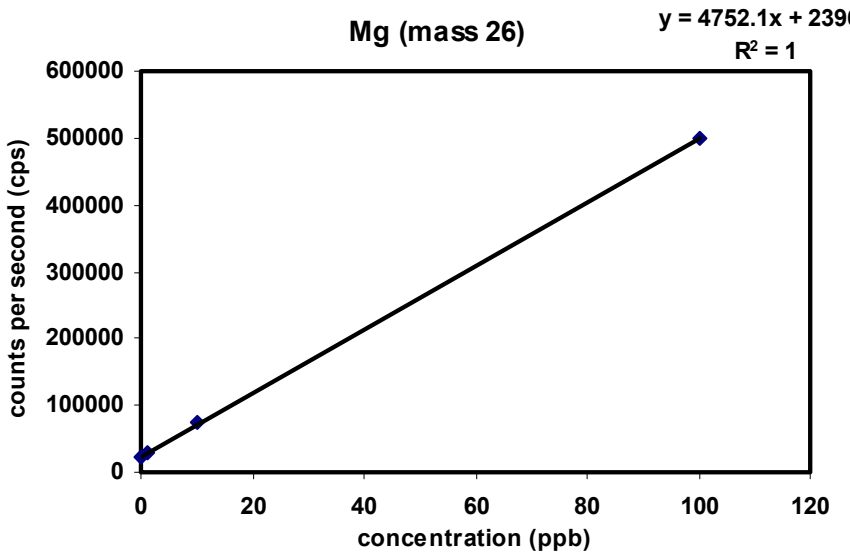
Appendix I (e)

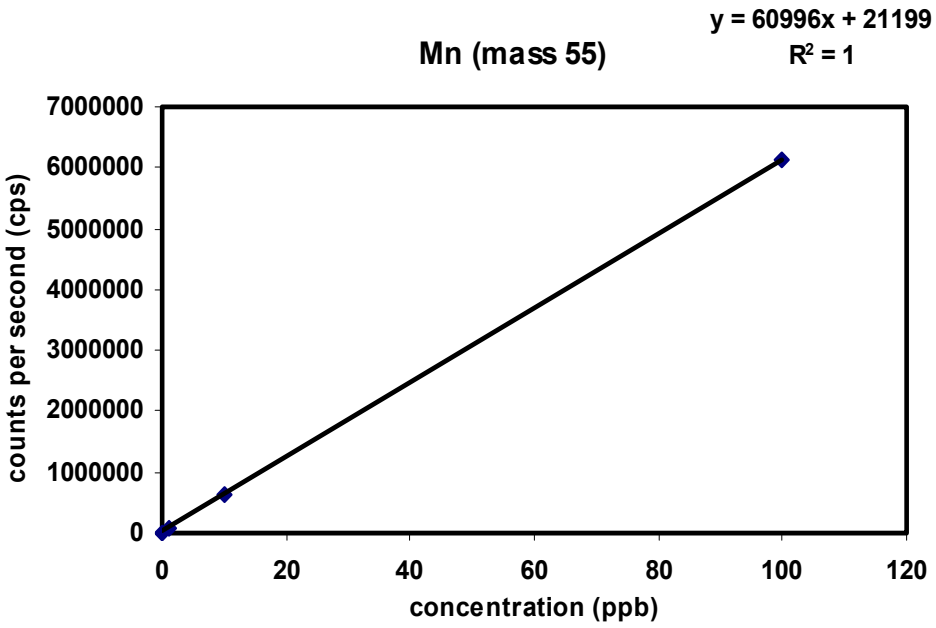
ICP-MS standard curves

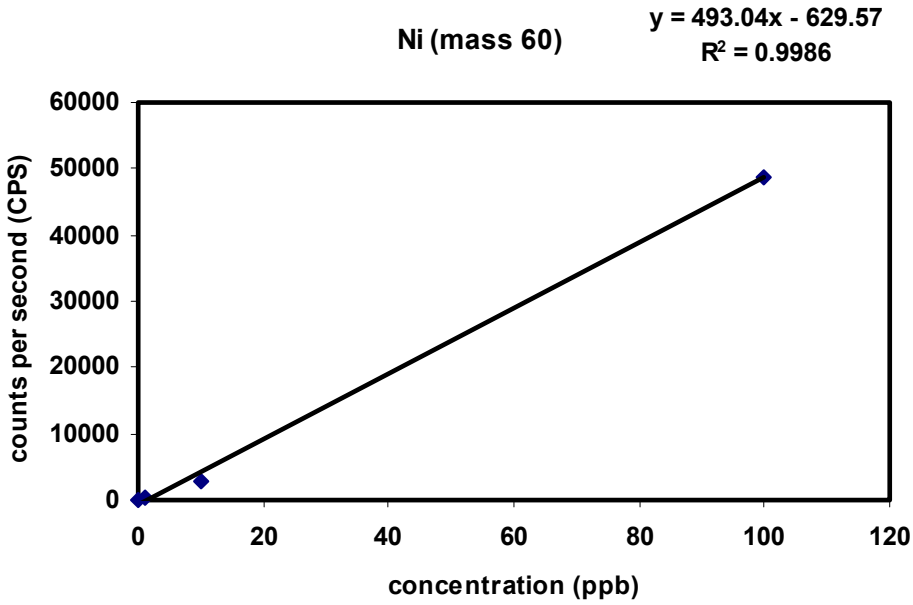


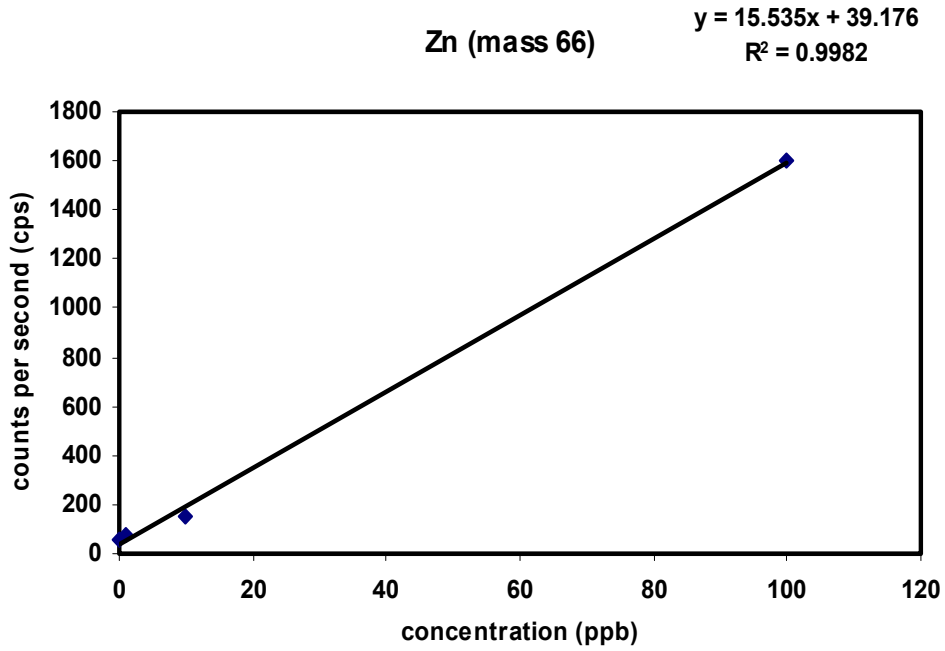



Fe (mass 56)
y = 470.35x + 672.15
R² = 1
counts per second (cps)
concentration (ppb)
0
10000
20000
30000
40000
50000
60000
0
20
40
60
80
100
120


Mg (mass 26)
y = 4752.1x + 23960
R² = 1
counts per second (cps)
concentration (ppb)
0
100000
200000
300000
400000
500000
600000
0
20
40
60
80
100
120


y = 60996x + 21199
R2 = 1
Mn (mass 55)
counts per second (cps)
7000000
6000000
5000000
4000000
3000000
2000000
1000000
0
0
20
40
60
80
100
120
concentration (ppb)

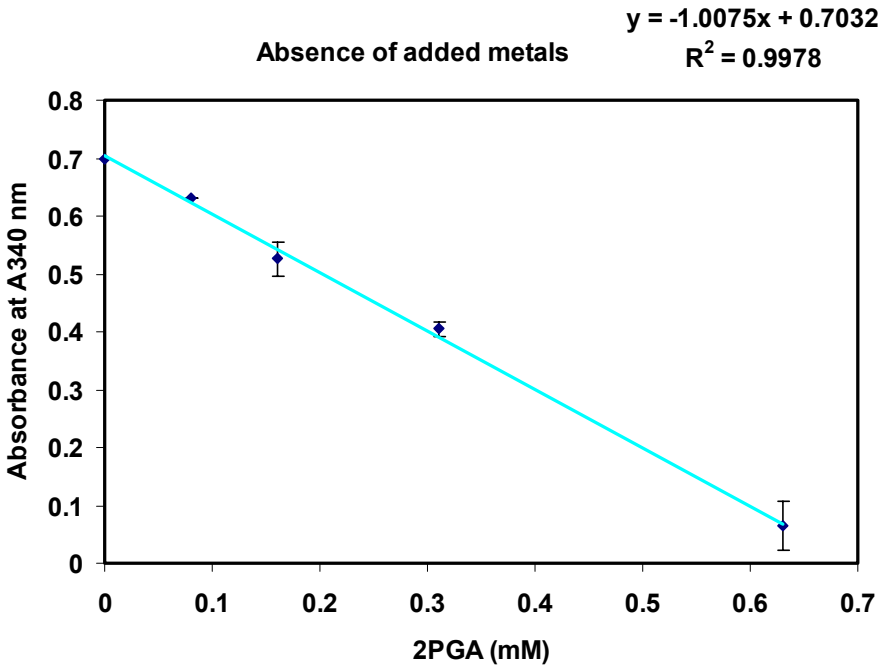

y = 493.04x - 629.57
R2 = 0.9986
Ni (mass 60)
counts per second (CPS)
60000
50000
40000
30000
20000
10000
0
0
20
40
60
80
100
120
concentration (ppb)

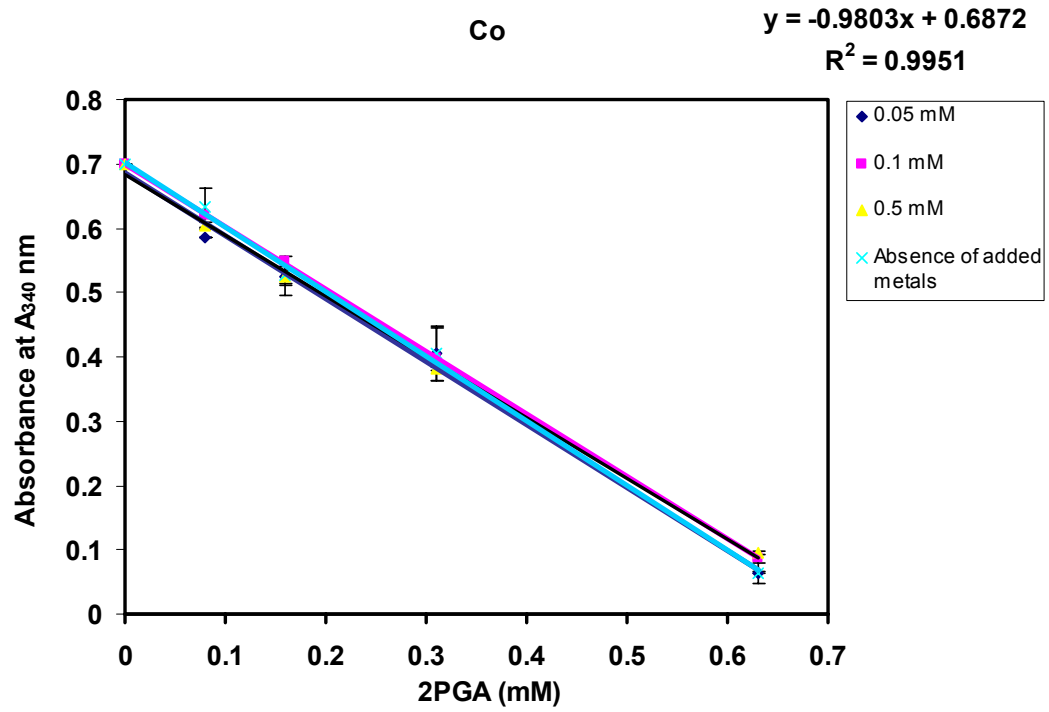


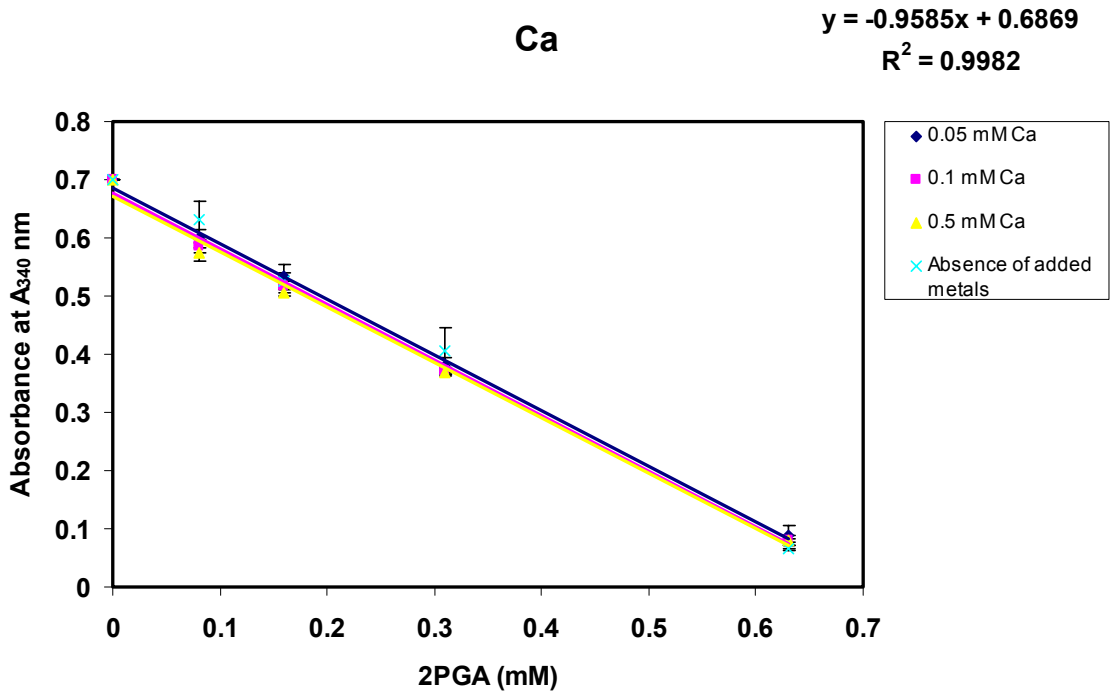
The calibration curves of concentration (in ppb) versus counts per second (cps) in ICP-MS analysis. The concentrations of divalent metals are in the range of 0, 1, 10 and 100 ppb (μg/L). The experiment was performed by using 20 mM TEA pH 7.6 and 50 mM NaCl.

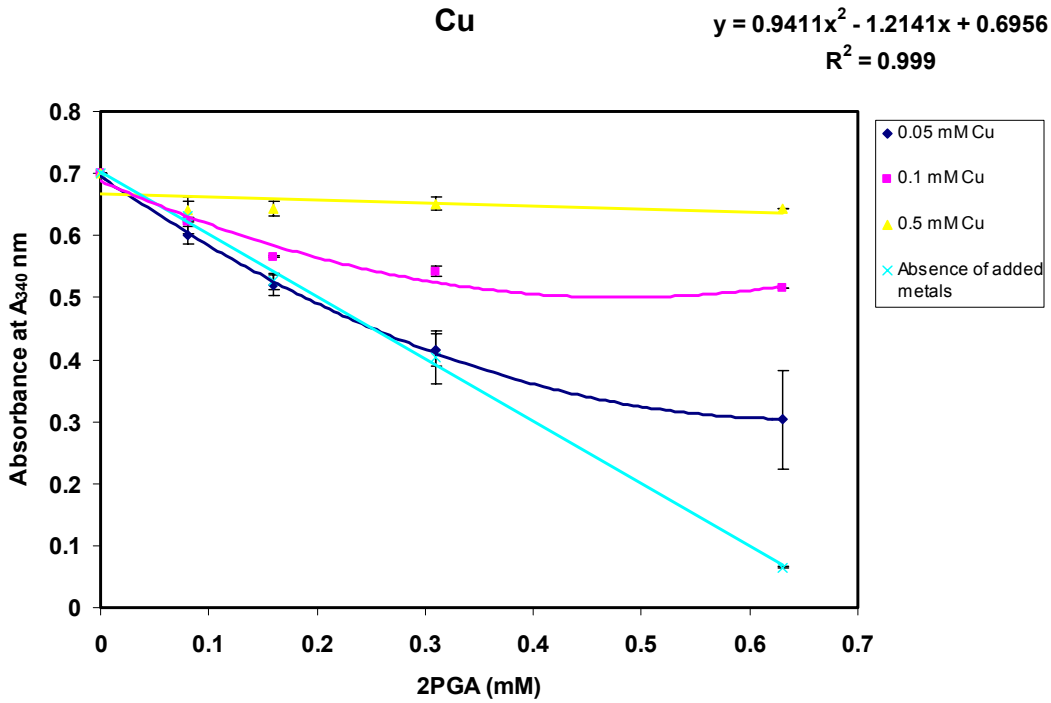
Appendix I (f)

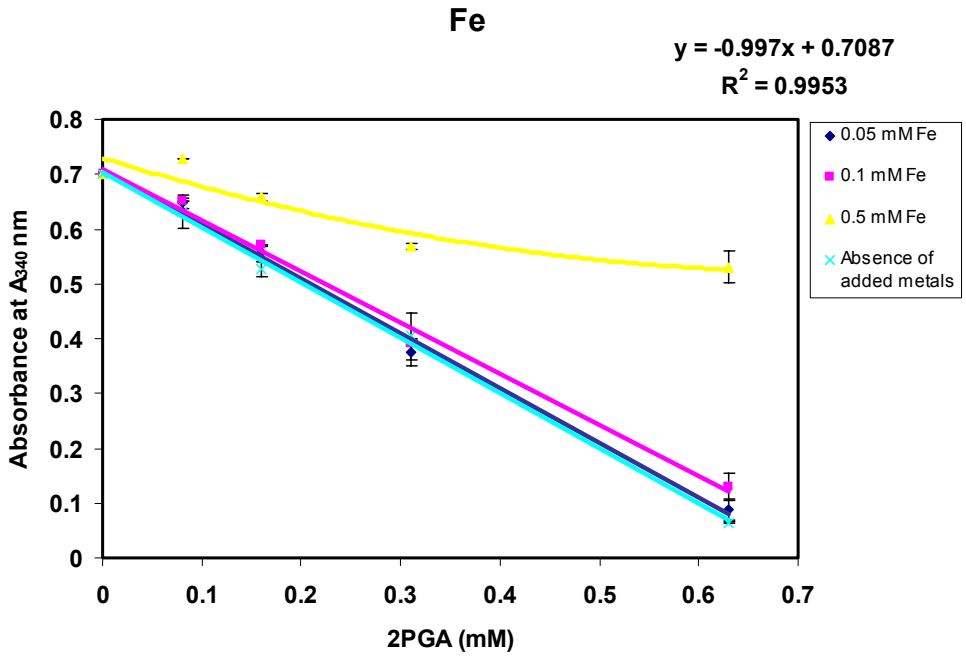
Discontinuous assay standard curves

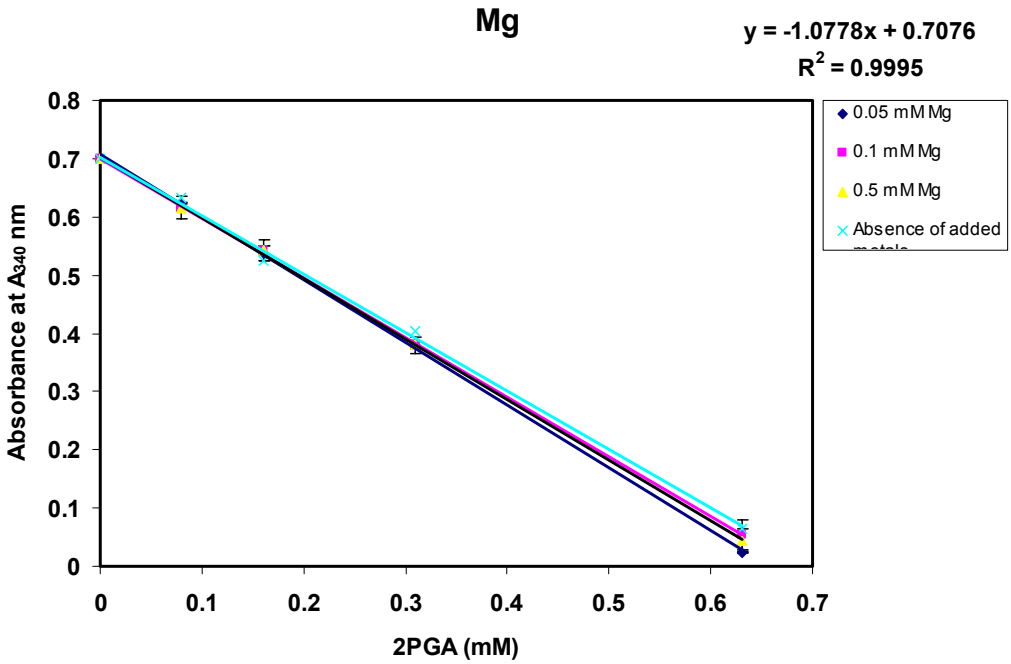


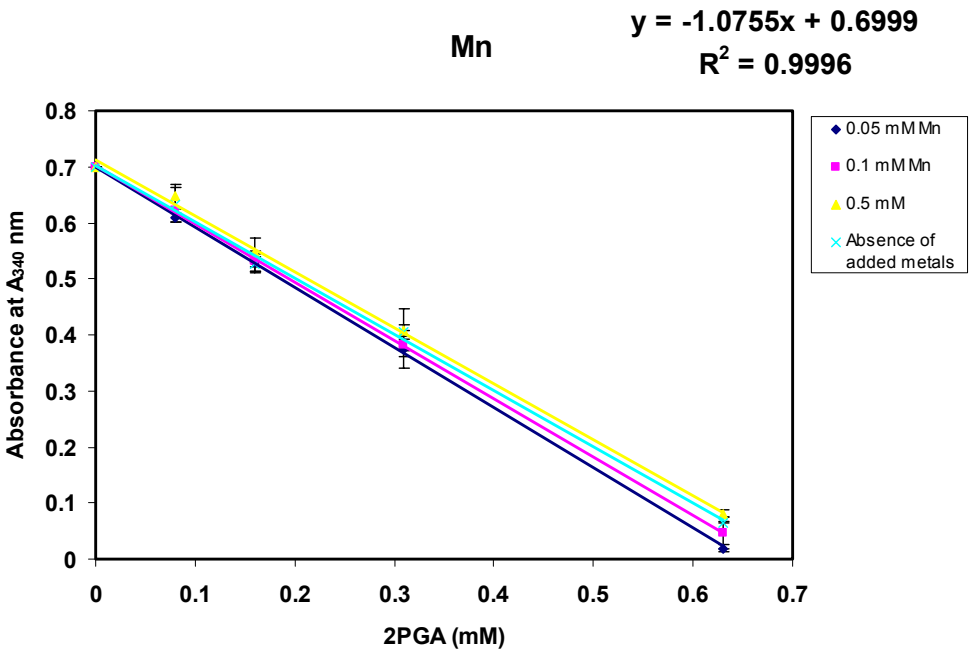


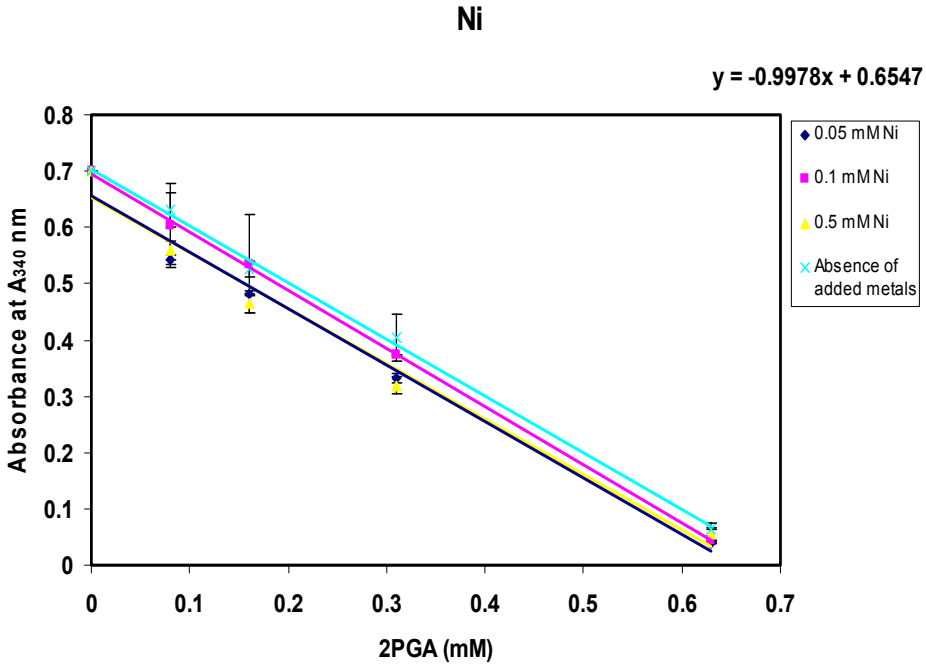

Ca
y = -0.9585x + 0.6869
R2 = 0.9982
Absorbance at A340 nm
2PGA (mM)
0.05 mM Ca
0.1 mM Ca
0.5 mM Ca
Absence of added metals

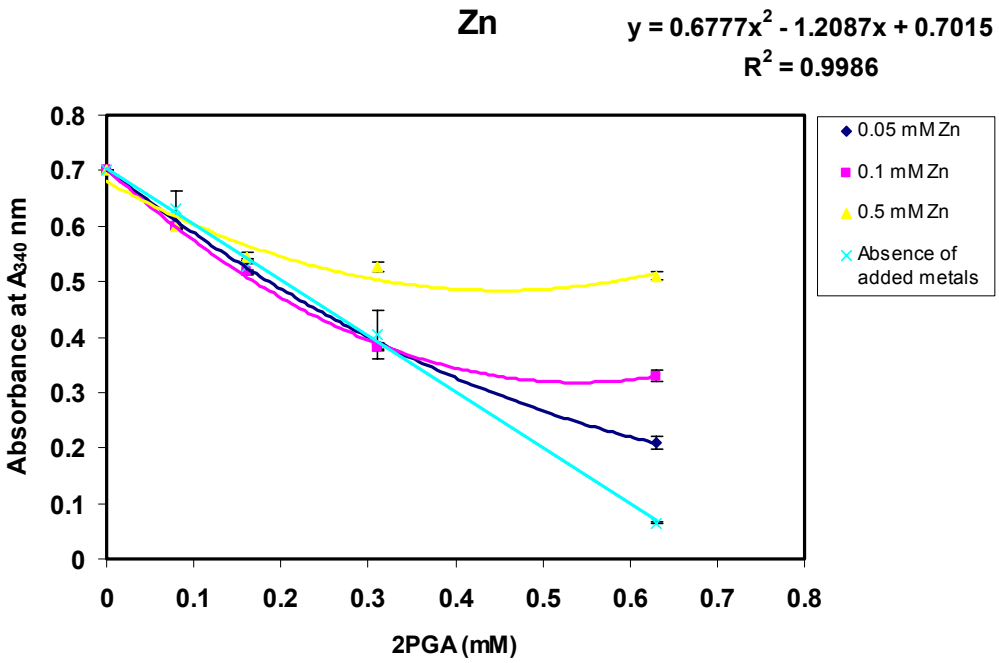

Cu
y = 0.9411x2 - 1.2141x + 0.6956
R2 = 0.999
Absorbance at A340 nm
2PGA (mM)
0.05 mM Cu
0.1 mM Cu
0.5 mM Cu
Absence of added metals


Fe
y = -0.997x + 0.7087
R2 = 0.9953
0.05 mM Fe
0.1 mM Fe
0.5 mM Fe
Absence of added metals
Absorbance at A340 nm
2PGA (mM)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8


Mg
y = -1.0778x + 0.7076
R2 = 0.9995
0.05 mM Mg
0.1 mM Mg
0.5 mM Mg
Absence of added
Absorbance at A340 nm
2PGA (mM)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8


Mn
y = -1.0755x + 0.6999
R2 = 0.9996
Absorbance at A340 nm
2PGA (mM)
0.05 mM Mn
0.1 mM Mn
0.5 mM
Absence of added metals


Ni
y = -0.9978x + 0.6547
Absorbance at A340 nm
2PGA (mM)
0.05 mM Ni
0.1 mM Ni
0.5 mM Ni
Absence of added metals



The discontinuous assay calibration standards were generated in the presence of three different concentrations of metals, and plotted as 2PGA concentrations versus NADH absorbance at A_{340} nm. The 2PGA concentrations which were used to start the reaction are in the range of 0.63 mM, 0.31 mM, 0.16 mM, 0.08 mM and 0 mM. The curve fit and R^2 value were only shown for the 0.05 mM concentration, as this is the calibration standard used for 2PGA concentration measurement.